METAL MINING GUIDANCE DOCUMENT FOR AQUATIC ENVIRONMENTAL EFFECTS MONITORING

JUNE 2002

TABLE OF CONTENTS

Chapter

1. INTRODUCTION........1

2. STUDY DESIGN........2

3. SITE CHARACTERIZATION........3

4. FISH MONITORING........4

5. BENTHIC INVERTEBRATE COMMUNITY MONITORING........5

6. EFFLUENT CHARACTERIZATION AND WATER QUALITY MONITORING........6

7. SUBLETHAL TOXICITY TESTING........7

8. SEDIMENT MONITORING........8

9. DATA ASSESSMENT AND INTERPRETATION........9

10. PUBLIC INVOLVEMENT........10

11. INFORMATION MANAGEMENT........11

12. ALTERNATIVE METHODS FOR BIOLOGICAL MONITORING........12

Table of contents

1. Introduction......1-1

1.1 Regulatory Framework......1-1

1.2 Program Overview......1-3

1.3 Part 1: Effluent and Water Quality Monitoring Studies......1-6
1.3.1 Effluent Characterization and Water Quality Monitoring......1-6
1.3.2 Sublethal Toxicity Testing......1-7
1.3.3 Effluent and Water Quality Monitoring Reports......1-7

1.4 Part 2: Biological Monitoring Studies......1-8
1.4.1 Study Design......1-8
1.4.2 Fish Survey......1-10
1.4.3 Benthic Invertebrate Community Survey......1-11
1.4.4 Alternative Monitoring Methods......1-11

1.5 Biological Monitoring Study Frequencies......1-12
1.5.1 First Biological Monitoring Study......1-12
1.5.1.1 Initial Monitoring......1-12
1.5.1.2 Mines Using Historical Information......1-12
1.5.2 Second and Subsequent Biological Monitoring Studies......1-13
1.5.2.1 Periodic Monitoring Studies......1-14
1.5.2.2 Focussed Monitoring Studies......1-15
1.5.2.3 Investigation of Cause Monitoring Studies......1-15
1.5.2.4 After Investigation of Cause......1-16
1.5.3 Conducting Biological Monitoring Studies......1-16

1.6 Data Assessment and Interpretation......1-16
1.6.1 Interpretative Reports for Biological Monitoring Studies......1-18
1.6.2 Interpretative Reports for Investigation of Cause Monitoring Studies......1-19

1.7 Recognized Closed Mines......1-19

1.8 Implementation of Environmental Effects Monitoring......1-20
1.8.1 Industry......1-20
1.8.2 Authorization Officer (AO)......1-20
1.8.3 Technical Advisory Panel (TAP)......1-20
1.8.4 Federal Government......1-21
1.8.5 Provincial/Territorial Governments......1-21

1.9 References......1-22

List of Figures

Figure 1-1: EEM program phases and activities........ 1-5
Figure 1-2: Sequence of activities and required timing for submission of the first Study Design and first Interpretative Report........ 1-13

103-04965/

1. Introduction

This Guidance Document provides information on the recommended methodologies, which are based on generally accepted standards of good scientific practice, and options on how to carry out the Environmental Effects Monitoring (EEM) studies. This document has been prepared based on recommendations from a series of multi-stakeholder Technical Subgroups (Fish, Benthic Invertebrate Community, Toxicity, Water and Sediment, and Stakeholder Involvement). Please note that this Guidance Document is published in a format and is viewed as a document which can, and will, be updated regularly as new information and research becomes available. To ensure that the most recent updates are included in this document, consult our website at www.ec.gc.ca/eem, or contact Environment Canada's, National EEM Office (phone: 819-997-1535; fax 819-953-0461; e-mail: eem-esee@ec.gc.ca).

The purpose of this introductory chapter is to provide the readers with a link between the terminology that is used in Section 7 and Schedule 5 of the MMER and what is used in this Guidance Document.

1.1 Regulatory Framework

The Assessment of the Aquatic Effects of Mining in Canada (AQUAMIN 1996) provided recommendations regarding the review and amendments of the *Metal Mining Liquid Effluent Regulations* (now called the *Metal Mining Effluent Regulations*, MMER) and the design of a national Environmental Effects Monitoring (EEM) program for metal mining. EEM is a requirement of the MMER under the authority of the *Fisheries Act*. The regulation requires all metal mines in Canada, to which the regulations apply, to conduct EEM *as per* the MMER.

The MMER prescribes limits for the discharge of deleterious substances, including arsenic, copper, total cyanide, lead, nickel, radium-226, zinc, pH of effluent, and total suspended solids (TSS), and a requirement for effluent to be non-acutely lethal to rainbow trout. These discharge limits were established to be minimum national standards based on best available technology economically achievable. The MMER defines "effluent" as mine water effluent, milling facility effluent, tailings impoundment area effluent, treatment pond effluent, treatment facility effluent – other than effluent from a sewage treatment facility - seepage and surface drainage that contains a deleterious substance.

Because the Canadian mining industry operates in many different types of ore deposits, uses a variety of mining and milling processes, and discharges effluents to a wide variety of receiving environments (see Chapter 3), the uniform effluent limits alone may not ensure adequate protection of all receiving environments. Thus, to evaluate the effects of mine effluent on the aquatic environment, site-specific EEM studies will be conducted by all mines subject to the MMER.

Many Canadian mines have previously conducted environmental monitoring in response to government priorities, emerging environmental issues and improved process and effluent treatment technologies (AQUAMIN 1996). The EEM program was designed to achieve national uniformity in monitoring of effects in the aquatic environment, while taking into consideration site-specific factors. Although this Guidance Document is not a legal document, it is intended to provide guidance on recommended standards of good scientific practice available to meet the EEM requirements.

Development of the EEM Program

In 1997, Environment Canada initiated a multi-stakeholder consultation to develop the EEM program requirements and detailed guidance for the metal mining EEM program, using the AQUAMIN recommendations as a starting point. Leading this process was a multi-stakeholder working group co-chaired by representatives of Environment Canada and the Mining Association of Canada. The EEM Working Group also included representatives of other federal departments, several provinces, the Cree Regional Authority and the Canadian Environmental Network. To assist in preparing these deliverables the Working Group established four subgroups to propose technical requirements and guidance in the following areas: fish monitoring; benthic invertebrate monitoring; effluent, water and sediment monitoring; and sublethal toxicity testing. Each of these subgroups was also multi-stakeholder, and like the working group, all subgroups worked to achieve consensus on issues related to both the requirements and the guidance.

During this same period, Natural Resources Canada lead an initiative called the Aquatic Effects Technology Evaluation (AETE). The objective of AETE was to evaluate environmental monitoring technologies to be used by the mining industry and regulatory agencies in assessing the impacts of mine effluents on the aquatic environment and to recommend specific methods or groups of methods that will permit accurate characterization of environmental impacts in the receiving waters in as cost-effective a manner as possible. AETE conducted detailed case studies at several mine sites across Canada, and also completed 20 detailed technical evaluations of a range of monitoring methods. The results of these case studies and technical evaluations, together with the final conclusions of AETE (AETE 1999) were given significant consideration by the EEM Working Group and the subgroups in the development of the proposed requirements and guidance for the EEM program. For more information on AETE, please go to the AETE web site: www.nrcan.gc.ca/mets/aete.

The consultation on the EEM program culminated in 1999 with a consensus agreement on proposed EEM requirements, and the completion of a guidance document which includes recommended monitoring methods for the fish survey, fish tissue and the benthic invertebrate community, and the testing of effluent, water and sediment. The proposed EEM requirements, as they appear in the MMER, is bases on this consensus document.

1.2 Program Overview

EEM is a scientific monitoring approach that can be used to help determine:

- effects in aquatic ecosystems caused by industrial effluent; and,
- the effectiveness of environmental protection measures.

The objective of metal mining EEM is to evaluate the effects of mine effluent on fish, fish habitat and the use of fisheries resources. This information may be used to contribute to an assessment of the adequacy of the MMER. A nationally consistent EEM program, along with social and economic information, can be used to assess the effectiveness of pollution prevention and control technologies, practices, programs, and indicate where there is a local, regional or national need for enhanced protection.

To meet this objective, the EEM program requires all mines regulated under the MMER to conduct ***biological monitoring studies*** *(Schedule 5, Part 2)*, which includes a fish survey (using indicators of fish population health and fish tissue analysis) and a benthic invertebrate community survey, and ***effluent and water quality monitoring studies*** *(Schedule 5, Part 1)* which includes effluent characterization, water quality monitoring and sublethal toxicity testing.

In meeting the objective of EEM, mines are encouraged to seek input from stakeholders, aboriginal communities, other interested regional parties and the public in recognition that the design, implementation and interpretation of the EEM study will benefit from their influence. Guidance and information on stakeholder involvement in the EEM process are provided in this document (Chapter 10).

EEM follows a tiered monitoring approach, with monitoring requirements of each phase dependent upon results of the previous monitoring phase. EEM studies will consist of both biological monitoring studies and effluent and water quality monitoring studies, but have the flexibility to take site-specific factors into consideration, as well as the results of prior studies at the site. Where there is more than one mine in close proximity, and effluents are discharged to the same drainage basin, joint EEM studies are encouraged.

For the purposes of this program, an "effect" is defined as follows *(Schedule 5, s. 1)*:

- an "effect on the fish population" means a statistical difference between fish population measurements taken in an exposure area and a reference area.
- an "effect on fish tissue" means measurements of total mercury that exceed 0.45 μg/g wet weight in fish tissue taken in the exposure area and that are statistically different from the measurements of total mercury taken in the reference area.

- an "effect on the benthic invertebrate community" means a statistical difference between benthic invertebrate community measurements taken in an exposure area and a reference area (e.g., control/impact design) or a statistical difference between measurements taken at sampling areas in the exposure area that indicate gradually decreasing effluent concentrations (e.g., a gradient design).

The guiding principles of the EEM program are that the program be scientifically defensible, cost-effective and provide flexibility for site-specific requirements, without subjecting field crews to unsafe sampling conditions. The program has also been designed to allow for incorporating new or improved monitoring techniques and to build on findings of relevant research programs or pilot studies.

A schematic overview of the EEM program phases, study activities and the related sequence of events is found in Figure 1.1.

There are four EEM study activities:

- submission of the study design for biological monitoring;
- conducting biological monitoring in the field;
- data assessment and interpretation, including submission of the interpretative report; and
- effluent and water quality monitoring and reporting.

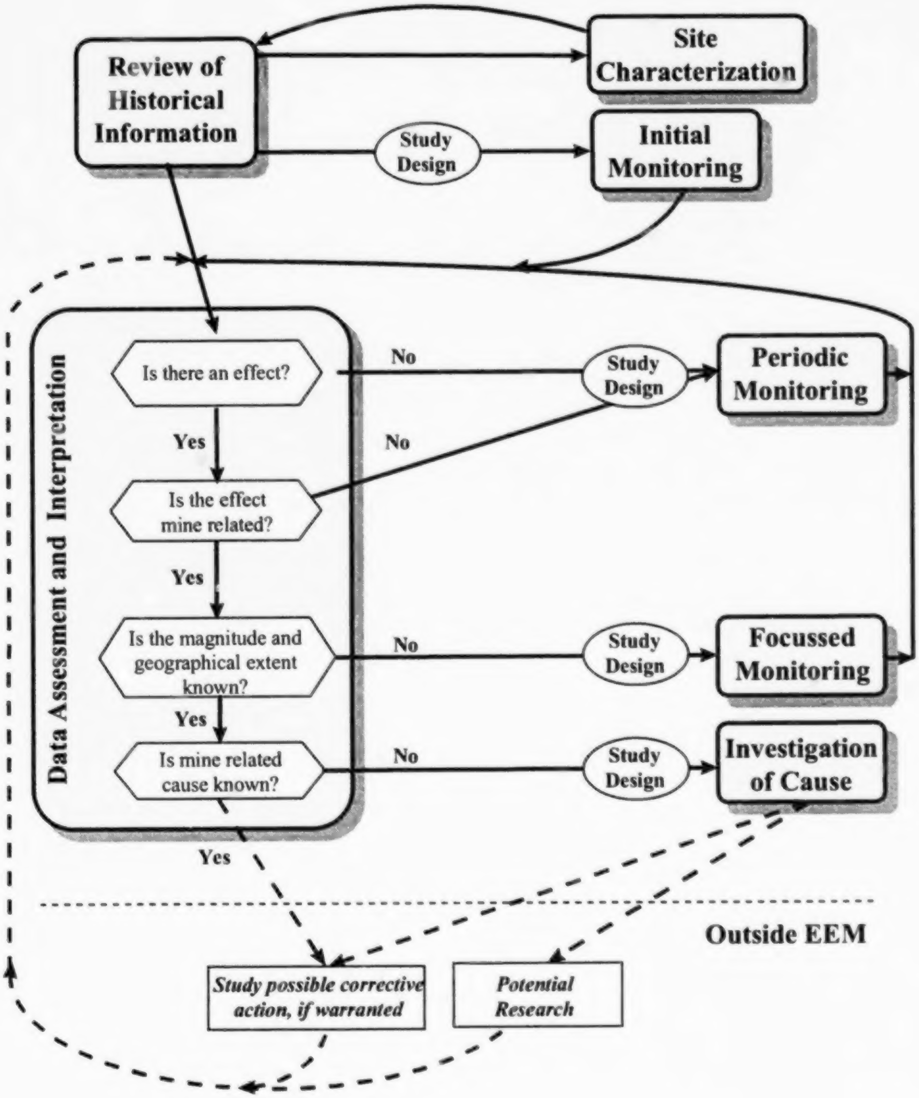
The proposed EEM program follows a tiered approach where results from previous biological monitoring studies are evaluated to trigger the next monitoring studies, in accordance with a pre-defined monitoring frequency.

The first biological monitoring study is called Initial Monitoring. Subsequent biological monitoring studies are dependent on whether or not effects are detected in fish, fish tissue and the benthic invertebrate community. These studies are called Periodic Monitoring, Focussed Monitoring or Investigation of Cause. Each of these phases have a specific objective and frequency, which is outlined in Section 1.5.

At individual sites, it is the responsibility of the mine operator to conduct EEM according to the requirements.

The proposed EEM requirements include two parts:
Part I: Effluent and Water Quality Monitoring Studies
Part II: Biological Monitoring Studies.

Figure 1-1: EEM program phases and activities


Review of Historical Information
Site Characterization
Study Design
Initial Monitoring
Data Assessment and Interpretation
Is there an effect?
No
Study Design
Periodic Monitoring
Yes
No
Is the effect mine related?
Yes
Is the magnitude and geographical extent known?
No
Study Design
Focussed Monitoring
Yes
Is mine related cause known?
No
Study Design
Investigation of Cause
Yes
Outside EEM
Study possible corrective action, if warranted
Potential Research

1.3 Part 1: Effluent and Water Quality Monitoring Studies

The purpose of effluent and water quality monitoring is to collect information about effluent and water quality, to provide background supporting information for the assessment and interpretation of the results of biological monitoring. Effluent and water quality monitoring consists of effluent characterization, sublethal toxicity testing of effluent, and receiving water quality monitoring.

Mines will begin effluent and water quality monitoring six months after they are subject to Section 7 of the MMER, and results will be submitted in a report to the authorization officer annually, by March 31.

1.3.1 Effluent Characterization and Water Quality Monitoring

Effluent characterization (Chapter 6) is conducted to provide information on the concentrations of compounds in effluent, in addition to the deleterious substances measured as part of the MMER, including other metals, nitrogen compounds, alkalinity and hardness *(Schedule 5, s. 4(1))*. Samples will be collected four times per year, no less than one month apart, on aliquots of effluent collected under Sections 12 and 13 of the MMER (which are the samples collected for the weekly effluent sampling to determine the pH and concentrations of the deleterious substances) *(Schedule 5, s. 4(2))*. It is recommended that, where possible, samples for effluent characterization are collected in each calendar quarter.

Water quality monitoring (Chapter 6) is conducted to provide information on the concentrations of contaminants in the exposure area as compared to the reference area. Samples will be collected four times per year, no less than one month apart, and will be analyzed for the same parameters as required for effluent characterization, as well as the deleterious substances and pH, temperature and dissolved oxygen *(Schedule 5, s. 7(2)(a))*. It is recommended that samples are collected in each calendar quarter, and at the same time are samples for effluent characterization are collected. Samples for water quality monitoring will be collected in both an exposure area and a reference area. These areas may not necessarily be the same as the sampling areas chosen for biological monitoring.

In addition, samples for water quality monitoring will also be collected in the sampling areas for biological monitoring, at the same time that biological monitoring is conducted *(Schedule 5, s. 7(2)(b))*.

1.3.2 Sublethal Toxicity Testing

Sublethal toxicity testing is conducted to determine if the effluent has the potential to affect fish, invertebrates, algae or plants and to track effluent quality. Endpoints measured in the sublethal toxicity tests include survival, growth and reproduction. Mines discharging effluent to a freshwater environment will be required to complete tests on a fish, an invertebrate, an algal and a plant species, while mines discharging effluent to an estuarine or marine environment will be required to complete tests on a fish, an invertebrate and an algal species *(Schedule 5, s. 5(1))*. Sublethal toxicity testing is required at the mine's final discharge point that has potentially the most adverse effect on the receiving environment *(Schedule 5, s. 5(2))*. More information on how to determine this is found in Chapter 7.

Sublethal toxicity testing is conducted using aliquots of effluent collected for effluent characterization. Sublethal toxicity testing will be required twice each calendar year for the first three years, beginning six months after the mine is subject to Section 7 of the MMER *(Schedule 5, s. 6(1))*. After the first three years, sublethal toxicity testing may then be reduced to once each calendar year.

If a mine has results of six sublethal toxicity tests that were conducted after December 31st, 1997, on a fish species, an invertebrate species and either a plant or an algal species, the mine may submit a report to the authorization officer indicating these results *(Schedule 5, s. 6(2))*. The mine may then conduct sublethal toxicity testing once each calendar year, beginning 6 months after the mine is subject to Section 7 of the MMER.

1.3.3 Effluent and Water Quality Monitoring Reports

All mines subject to the MMER will submit to the authorization officer an Effluent and Water Quality Monitoring Report no later than March 31 each year, which includes the following information, with respect to effluent and water monitoring completed in the previous calendar year *(Schedule 5, s. 8)*:

- dates when all samples were collected for sublethal toxicity testing, effluent characterization and water quality monitoring;

- location(s) of final point(s) of discharge from which samples were collected for sublethal toxicity testing and effluent characterization;

- latitude and longitude of sampling areas for water quality monitoring, in degrees, minutes and seconds and written descriptions that can identify the location of the sampling areas;
- results of effluent characterization, water quality monitoring and sublethal toxicity testing;

- the methodologies used to conduct the analyses for effluent characterization and water quality monitoring and the related detection limits; and,

- a description of quality assurance and quality control measures that were implemented and the data related to the implementation of these measures.

Other information in the effluent and water quality monitoring report may include:

- for any sublethal toxicity test that has an IC25 result of less than 30%, the spatial extent in the exposed area where concentrations of effluent are comparable to the IC25 results;

- estimated mine related changes in measurements in the exposure area of MMER deleterious substances, and parameters measured; and,

- comparison of water quality data in exposure and reference areas.

1.4 Part 2: Biological Monitoring Studies

1.4.1 Study Design

The objective of a study design is to outline what field work and associated laboratory work will be done to complete the biological monitoring portion of the EEM study. Achieving an appropriate study design for each mine will be an iterative process.

Prior to any biological monitoring studies being conducted, a study design will be submitted to the authorization officer at least six months prior to the commencement of conducting the biological monitoring in the field *(Schedule 5, s. 15)*. This will allow reasonable time for the Authorization Officers to review and provide comments to the mine regarding the study design.

The study design will include *(Schedule 5, s. 10)*:

- a detailed timetable for conducting each of the biological monitoring study requirements;
- a description of site characterization information (Chapter 3);
- a description of how the fish survey will be conducted in order to determine if there are effects on fish population and fish usability (tissue analysis) (Chapter 4);
- a description of how the benthic invertebrate community survey will be conducted in order to determine if there are effects on fish habitat (Chapter 5);

- identification of the quality assurance and quality control measures that will be taken to ensure the validity of the data; and;
- a summary of any previous biological monitoring conducted for EEM.

Other recommended details of the study design may include:

- the goals and expectations of the EEM study;
- the overall approach, including rationale based on site characterization information and previous monitoring results;
- statistical design: development of hypotheses, selection of statistical methods, determination of sample size (statistical significance and power analysis);
- the design of the monitoring study, taking into account confounding factors (if any): where to sample, what to measure, how frequently to sample;
- operating plans and procedures: sampling procedures, laboratory analysis procedures, quality assurance and quality control procedures, data storage and retrieval, data analysis;
- a plan for data interpretation and evaluation.

The concepts and basic elements of a good measurement system, quality assurance/ quality control (QA/QC) are discussed in this Guidance Document for each component. QA/QC procedures should be set *a priori* as part of the study design, and those that will be implemented should be described, in order to ensure validity of the data. Quality assurance results that may affect the reliability of the conclusions will be submitted as part of each interpretative report.

Site Characterization

Each biological monitoring study design will contain site characterization information (Chapter 3), which is information that is needed to prepare an EEM study design. Site characterization information will be updated (if necessary) and used in designing each study at different phases of the EEM program. There is a variety of information about the mine operation, and the exposure and reference areas, which is essential for the preparation of an EEM study design (see Chapter 2). As well, information that is needed includes *(Schedule 5, s. 11)*: mine operation information and a description of the study area, including geological, oceanographical, limnological, hydrological, chemical and biological features of the area. A description of how the effluent mixes within the exposure area will also be included.

A description of confounding factors will be included as part of site characterization. The term "confounding influence or factor" refers to any anthropogenic, natural or other factors, that are not related to the effluent under study but may reasonably be expected to contribute to an observed effect. Mines are not required to monitor effects that are caused by confounding influences. However, where confounding influences are present in the study area, mines will:

- identify and describe potential confounding influences as part of site characterization within the study design
- where possible, design EEM studies to resolve the difficulties related to data interpretation caused by confounding influences.

1.4.2 Fish Survey

The fish survey (Chapter 4) will consist of a study of the ***fish population*** (by conducting a fish survey using indicators of fish population health) to determine if mine effluent is having an effect on the fish, and a study of ***fish tissue*** which will determine if effluent is having an effect on fish usability.

Note that a mine is not required to conduct a fish study if the concentration of the effluent in the exposure area is less than 1% within 250m of each final discharge point *(Schedule 5, s. 9(b))*. The description of how the effluent mixes with the receiving water is included in the study design as part of site characterization. A fish tissue analysis is only required if the mine has measured concentrations of mercury greater than or equal to 0.10 µg/L in the effluent during effluent characterization *(Schedule 5, s. 9(c))*.

The fish survey is conducted to determine if there have been changes in fish growth, reproduction condition and survival. The scientifically defensible method recommended to determine if there are changes in fish growth, condition, reproduction and survival is by collecting fish species found in the exposure and reference area, and comparing measurements of length, weight, gonad size, liver size, fecundity, and egg size. Results for fish collected in the exposure area will be compared statistically with those from fish collected in the reference area.

The fish tissue study will be conducted if mercury is present in the effluent at concentrations equal to or greater than 0.10 µg/L, as determined in effluent characterization. It is recommended that tissue samples from a fish species that is locally consumed will be collected from a minimum of eight fish. Mercury concentrations in tissue samples from the exposure area will be compared with the mercury concentrations in samples from the reference area, and with the Health Canada guideline of 0.45 µg/g wet weight for mercury in fish tissue.

1.4.3 Benthic Invertebrate Community Survey

Mines will be required to conduct a benthic invertebrate community survey (Chapter 5) to determine if their effluent is having an effect on fish habitat. This is done by collecting benthic invertebrates in the exposure area and reference area or a gradient of exposure and comparing measurements of benthic invertebrate density, taxa richness, the Simpson's Diversity Index and the Bray-Curtis Index.

To ensure site-specific flexibility mines can choose from any of several scientifically defensible sampling designs, including control/impact, gradient, and the reference condition approach. Sample sizes (i.e., number of sampling stations) are determined site-specifically using statistical power analysis. Samples from both the exposure and reference areas will be collected in the most "ecologically relevant" area (i.e. considering habitat type with the highest benthic invertebrate diversity and the dominant habitat in the exposure area) and season (i.e. the time of year when the benthic invertebrate diversity is highest and benthic invertebrates are most exposed to effluent).

The mine will report the total organic carbon content and particle size distribution of the sediment (Chapter 8) collected at each benthic invertebrate sampling station.

1.4.4 Alternative Monitoring Methods

At some mine sites the biological monitoring methods described above (a wild fish survey and benthic invertebrate community survey) may not be appropriate. It should be emphasized that alternative monitoring methods should only be considered when field surveys cannot be conducted. See Section 12.1.2.2. for the applicability of alternatives in the metal mining EEM program. The reasons for this are site specific, but the most common reasons are the presence of hazardous conditions (e.g. high velocity) or the presence of confounding factors such as other effluent discharges in the exposure area, that will make it difficult or impossible to isolate any effects attributable to the effluent being monitored.

Where mines cannot design the fish or benthic invertebrate community surveys to resolve difficulties associated with confounding influences, mines will provide a scientific rationale and justification and propose cost effective and technically feasible alternative monitoring methods within the study design. Chapters 4 and 5 describe potential alternative biological monitoring methods

A number of alternative monitoring methods are recommended in this Guidance Document (Chapter 12). Mines may choose these or other scientifically defensible monitoring methods provided that the results can determine if the effluent is having effects on the fish population (growth, reproduction, condition and survival), fish tissue (mercury levels) or the benthic invertebrate community (benthic invertebrate density, taxa richness, the Simpson's Diversity Index and the Bray-Curtis Index). Recommended alternative monitoring methods to the fish survey are mesocosm studies, caged bivalves, and expanded sublethal toxicity testing. For benthic invertebrate community surveys, the recommended alternative monitoring method is mesocosm studies.

1.5 Biological Monitoring Study Frequencies

1.5.1 First Biological Monitoring Study

1.5.1.1 Initial Monitoring

The first biological monitoring study for a mine that does not have any previous bioloigcal monitoring data (that can determine if the effluent is having an effect on the fish population, fish tissue and the benthic invertebrate community) is an ***Initial Monitoring*** study. The study design will contain the information outlined in Section 1.4.1 and be submitted within 12 months of being subject to Section 7 of the MMER *(Schedule 5, s. 14(a))*. The Interpretative Report will then be submitted within 30 months of the mine being subject to Section 7 of the MMER (Figure 1.2*) (Schedule 5, s. 18(a))*.

The Initial Monitoring studies will be designed to determine if the effluent is causing an effect, by conducting biological monitoring studies, when required, on:

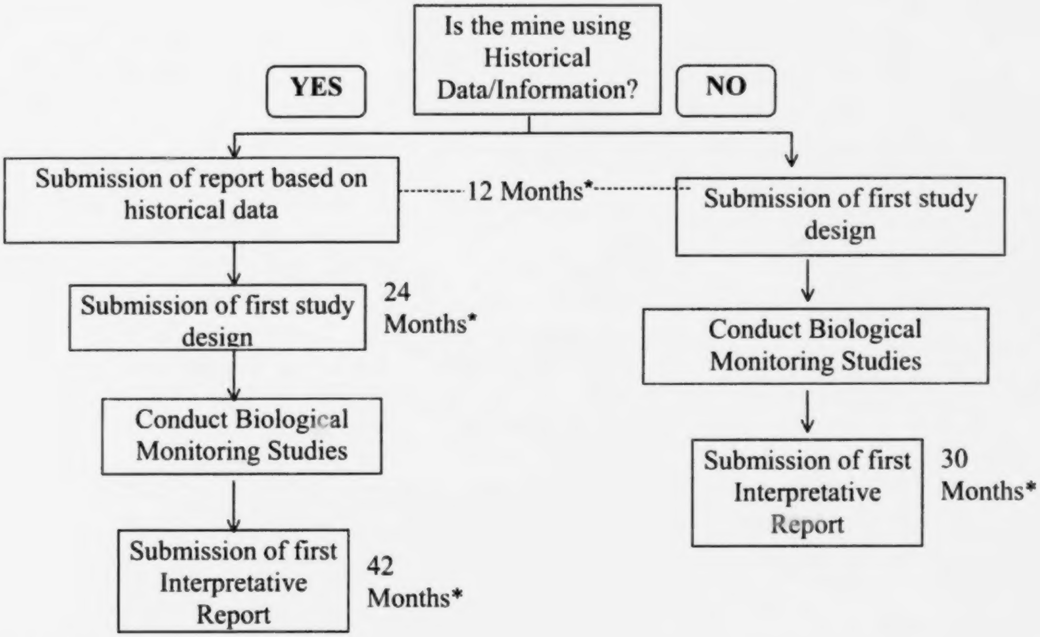
- fish;
- fish tissue; and
- the benthic invertebrate community.

1.5.1.2 Mines Using Historical Information

Many mines may have collected biological monitoring information prior to the coming into force of the MMER, which may be used as part of the EEM program. If there is previous biological monitoring information that can determine if there are effects on fish, fish tissue or the benthic invertebrate community, the mine may submit a report to the authorization officer, that includes the results of the historical biological monitoring studies, within 12 months of being subject to Section 7 of the MMER (Figure 1.2) *(Schedule 5, s. 14(b))*.

In such a case, after submission of the report, the mine will submit a study design within 24 months of being subject to Section 7 of the MMER. Depending on how much data a mine has, the purpose of this study design would be to design the first EEM study to confirm previous monitoring results (Periodic Monitoring), begin to determine the magnitude and geographic extent of the effect (Focussed Monitoring), or to investigate the cause of effects (Investigation of Cause). This first study design will also answer questions that have not been previously examined on either fish, fish tissue or benthic invertebrate community. The study design will be submitted no later than 6 months prior to conducting the biological monitoring. The Interpretative Report will then be submitted within 42 months of the mine being subject to EEM requirements (Section 7) of the MMER *(Schedule 5, s. 18 (b))*.

Figure 1-2: Sequence of activities and required timing for submission of the first Study Design and first Interpretative Report


Is the mine using Historical Data/Information?
YES
NO
Submission of report based on historical data
12 Months*
Submission of first study design
Submission of first study design
24 Months*
Conduct Biological Monitoring Studies
Conduct Biological Monitoring Studies
Submission of first Interpretative Report
30 Months*
Submission of first Interpretative Report
42 Months*

* Indicates number of months from the date the mine becomes subject to Section 7 of the MMER.

1.5.2 Second and Subsequent Biological Monitoring Studies

The frequency of monitoring for the second and any subsequent biological monitoring studies is dependent on previous results. Study designs will be submitted at least 6 months before each biological monitoring study is conducted.

Study designs for the second and subsequent biological monitoring studies include the following information:

- information outlined in Section 1.4.1
- a summary of the results of previous biological monitoring studies including:
    - the fish survey
    - fish tissue analysis
    - the benthic invertebrate community survey including sediment characterization data (particle size and total organic carbon)
- a description of the additional sampling areas if the next EEM phase conducted is Focussed Monitoring (Section 1.5.2.2)
- a description of the field and laboratory methods that will be used if Investigation of Cause is conducted (Section 1.5.2.3).

It is understood that a mine may proceed to a Focussed Monitoring study or an Investigation of Cause monitoring study, or corrective actions at any time in the EEM process.

1.5.2.1 Periodic Monitoring Studies

The general rule for the submission of the second and subsequent EEM interpretative report is 36 months of the preceeding interpretative report *(Schedule 5, s. 22(1))*. This is called ***Periodic Monitoring - Surveillance.*** Surveillance monitoring is conducted when:

1. no effects have been found in fish, fish tissue ***and*** the benthic invertebrate community in one study

2. an effect has been found in one or two of the fish, fish tissue or benthic invertebrate community (but not all three) in one study or,

3. the magnitude and geographic extent and cause of the effects are known.

When an **effect** has been found in fish, fish tissue ***and*** the benthic invertebrate community, the mine will submit its next interpretative report within 24 months *(Schedule 5, s. 22(2)(a))*. This is called ***Periodic Monitoring - Confirmation***. The purpose of the confirmation phase is to confirm if effects found in previous monitoring studies are reproducible and related to the mine. Mines are encouraged to re-evaluate their study design prior to conducting confirmation monitoring, especially if it is questionable whether the effect found is mine related. A confirmed difference (e.g., same effect found in two consecutive monitoring phases) between a reference and exposure area will be assumed to be a mine related effect.

If, after **two** consecutive surveys, **no effects** have been found in fish, fish tissue ***and*** the benthic invertebrate community, the mine may submit the next interpretative report within 72 months *(Schedule 5, s. 22(2)(b))*. This is ***Periodic Monitoring - Minimal.*** If there are changes in the mine operations or environmental conditions that result in an increased potential to cause an effect in the aquatic environment, it is recommended that confirmation monitoring of fish, fish tissue and the benthic invertebrate community be conducted within 24 months of the change.

1.5.2.2 Focussed Monitoring Studies

If **two** consecutive preceding EEM studies found the same **effect** on fish, fish tissue ***or*** the benthic invertebrate community the mine is required to submit their next interpretative report within 24 months of the previous report *(Schedule 5, s. 22(2)(c))*. This is called ***Focussed Monitoring***. The purpose of Focussed Monitoring is to determine the magnitude and geographical extent of the effect.

The design of Focussed Monitoring studies will be site-specific and will require sampling additional exposure areas progressively further from the effluent discharge, until the effect is no longer present. The scope may be broadened to include other environmental components, additional indicators or measurements to help define the magnitude of effect and clarify the understanding of effects.

Focussed monitoring studies may address the following questions:

- Are other components of the aquatic ecosystem likely to be affected and to require monitoring?
- Is the effect of an acute or chronic nature?
- Is the effect related to chemical contamination or physical disruption, or both?
- Are fish and/or the benthic invertebrate community directly affected or are effects mediated through their food webs?

1.5.2.3 Investigation of Cause Monitoring Studies

If previous Focussed Monitoring studies have determined the magnitude and geographic extent of an effect and the cause of the effect is still not know, the mine will conduct an ***Investigation of Cause*** monitoring study to determine the cause of the effect. The Investigation of Cause interpretative report will be submitted within 24 months of the previous interpretative report *(Schedule 5, s. 22(2)(c))*.

The questions that should be considered during Investigation of Cause monitoring studies are:

- What are the specific causes of the effect under investigation?
- What components of the effluent are responsible for the effects?

1.5.2.4 After Investigation of Cause

Once the cause of the effect is known, the mine will conduct Periodic Monitoring - surveillance, to ensure effects are not changing.

The mine may also proceed to study possible corrective actions or potential research work, which is outside of the EEM program. Environment Canada, in consultation with Fisheries and Oceans Canada, has developed a "Guidance for Determining Follow-up Actions when Effects Have Been Identified in Environmental Effects Monitoring (EEM)". This document provides guidance to regulatory agencies, regulated facilities, and EEM practitioners for determining follow-up actions when effects have been identified in EEM. It discusses the factors that need to be considered for determining follow-up actions, and explains the roles and responsibilities of government, industry, and stakeholders in this process.

1.5.3 Conducting Biological Monitoring Studies

Biological monitoring studies are conducted according to the study design that was submitted. It is understood that circumstances may arise that make it impossible to follow the study design. The mine will inform the authorization officer as soon as possible of the changes, and these changes will be documented in the interpretative report.

1.6 Data Assessment and Interpretation

Following biological monitoring, data assessment and interpretation will be conducted to determine if mine effluent is causing an effect and what future monitoring requirements will be. Chapter 9 describes the specific data assessments that are done in order to determine if there are effects on fish, fish tissue or benthic invertebrate community.
Four key questions that may be addressed, depending on the monitoring phase, in data assessment and interpretation are:

- Is there an effect?
- Is the effect mine related?
- Is the magnitude and geographic extent of the effect known?
- Is the mine related cause of the effect known?

Is there an effect?

An effect is defined as an effect on fish or an effect on the benthic invertebrate community, where:

- an "effect on the fish population" means a statistical difference between fish population measurements taken in an exposure area and a reference area.
- an "effect on fish tissue" means measurements of total mercury that exceed 0.45 µg/g wet weight in fish tissue taken in an exposure area and that are statistically different from the measurements of total mercury in fish tissue in a reference area.
- an "effect on the benthic invertebrate community" means a statistical difference between benthic invertebrate community measurements taken in an exposure area and a reference area (e.g., control/impact design) or a statistical difference between measurements taken within the exposure area at stations that indicate gradually decreasing effluent concentrations (e.g., a gradient design).

Is the effect mine related?

When an effect is observed during Initial or Periodic Monitoring studies, the mine will determine if the effect is mine related. The effect is assumed not to be mine related until confirmation of the effect is completed during a Periodic Monitoring study. When an effect previously seen is confirmed the mine will proceed to the next question of data assessment and interpretation. If the effect is not mine related, the mine will conduct a Periodic Monitoring (surveillance) study and continue monitoring to ensure that the effects are not mine related, the effects are not changing or that there are no new effects.

Is the magnitude and geographical extent of the effect known?

When a mine related effect is found, the mine is required to determine its magnitude and extent by conducting a Focussed Monitoring study. If the magnitude and extent of the effect are known, then the mine will proceed to the next question.

Is there a mine related cause of the effect known?

When a mine related effect on the fish, fish tissue or the benthic invertebrate community is observed, and its extent and magnitude is known but previous monitoring has failed to provide a satisfactory explanation of the cause of the effect, the mine is required to conduct an Investigation of Cause monitoring study.

Once a cause of the effect has been identified, the corrective actions may be considered. This is outside EEM. When the mine related cause is know, the mine continues monitoring the effect (Periodic Monitoring - Surveillance) to ensure there are no new effects and the effects are not changing.

1.6.1 Interpretative Reports for Biological Monitoring Studies

The frequency for submission of interpretative reports to the authorization officer can be found in Schedule 5 of the MMER. Interpretative reports for the first and for subsequent biological monitoring studies completed as part of the EEM program will include the following information:

1) A description of any deviations from the study design that occurred in meeting the requirements, including:
   - the impact of those deviations on the EEM study; and,
   - the adaptations made to the study as a result of the deviations.

2) Location of sampling areas, including the latitude and longitude, as well as a written description of the sampling area.

3) Dates and times when the samples were collected.

4) Sample sizes.

5) A summary of the results, including the arithmetic mean, the standard error, the standard deviation, the minimum and maximum values for each of the benthic invertebrate community sampling areas:
   - benthic invertebrate density;
   - taxa richness;
   - Simpson's diversity index; and,
   - Bray-Curtis index.

   It is recommended that an appendix is included in the interpretative report, outlining all measurements recorded, compiled for the following headings:
   - fish survey;
   - fish tissue; and,
   - benthic invertebrate community.

6) A summary of the results, including the arithmetic mean, the standard error, the standard deviation, the minimum and maximum values, for the fish measurements (growth, reproduction, condition and survival) and fish tissue analysis for samples taken in the exposure and reference areas:

7) The results and statistical power of the statistical analyses, conducted on data reported according to 3 and 4 above, compiled for the following headings:
   - fish population survey;
   - fish tissue analysis;
   - benthic invertebrate community survey.

8) Identification, based on the scientific statistical analyses, of:
    - the presence or absence of an effect; or,
    - the magnitude and geographic extent of the effect.

9) A comparison of the results of the statistical analysis conducted and sublethal toxicity testing reported since the preceding interpretative report, to determine if there is a correlation.

10) The conclusions to the EEM study, and identification of how the conclusions will impact the study design for the subsequent EEM study.

In preparing the conclusions, the following factors will be taken into consideration:

- results of previous EEM studies;
- the presence of confounding influences;
- the results of the statistical power analysis; and,
- any quality assurance or quality control results that may interfere with the reliability of the conclusions.

11) A schedule outlining:
    - when the next study design will be submitted;
    - when the next monitoring will be conducted; and
    - when the next interpretative report will be submitted.

12) Any recommendations for improvements or changes to the study design to be developed for the subsequent EEM study.

1.6.2 Interpretative Reports for Investigation of Cause Monitoring Studies

Interpretative reports submitted when the mine is conducting an Investigation of Cause study may not contain all of the same information as other biological monitoring studies. The reports should include a description of any deviations from the study design that was submitted, the dates and times when the samples were collected, a description of how the conclusions will impact subsequent study designs, and the dates and times when the next monitoring will occur. The interpretative reports will include information pertaining to the cause of the effect that is being studied.

1.7 Recognized Closed Mines

An owner or operator of a mine can become identified as a recognized closed mine if the mine's rate of production is less than 25% of its design rated capacity for a period of three years, starting the day the mine provides written notice to the AO of the mines intent. In order to fulfil part of the requirements to become a recognized closed mine, the mine will submit to the AO a biological monitoring study design within six months of providing the written notice. The mine will base the monitoring phase (Periodic Monitoring, Focussed Monitoring, or Investigation of Cause) on the results of the previous biological monitoring study. The study will be conducted based on the study design. The interpretative report will be submitted to the authorization officer within 36 months of the day the written notice was provided by the mine to the authorization officer. Effluent characterization and water quality monitoring will continue until the mine obtains recognized closed mine status.

1.8 Implementation of Environmental Effects Monitoring

1.8.1 Industry

Each owner or operator of a mine regulated under the *Metal Mining Effluent Regulations* is responsible for ensuring that EEM is completed in accordance with the MMER.

1.8.2 Authorization Officer (AO)

The AO is the Regional Director of Environmental Protection, Environment Canada as prescribed in the MMER (Schedule 1) . The Authorization Officer will:

- form Technical Advisory Panel(s) (TAP), made up of representatives which may include representatives from Environment Canada or other Federal departments; representative(s) of the relevant province(s) will be invited to participate; other stakeholders may also be invited to participate; the panel will be chaired by the AO or his/her designate;
- review and provide comments on EEM study designs, based on recommendations of the TAP;
- review and provide comments on EEM interpretative reports, based on recommendations of the TAP;

1.8.3 Technical Advisory Panel (TAP)

The responsibility of the TAP is to ensure the proper design and implementation of EEM proposed by the mines and to ensure study results meet the objectives of the EEM program. To this end the TAP will:

- advise the AO in discussions with the mines;
- make recommendations to the AO on adequacy of study designs (e.g. level of resolution, sampling station location, quality assurance/quality control (QA/QC) practices);
- review the interpretative report submitted by the mines and advise the AO as to:
  - the quality of data and any required modification to QA/QC procedures;
  - the validity of conclusions presented in the report
  - the frequency of subsequent monitoring study, based on the previous monitoring results.

1.8.4 Federal Government

The information provided by the EEM Program will in most instances enable the federal government to determine the extent to which mine effluents are affecting the aquatic environment, particularly with respect to the protection of fish, fish habitat and the use of fisheries resources. The EEM data will determine if there is a statistically significant effect, and will determine the magnitude and geographic extent of the effect. This information will be considered by federal regulators to determine the effectiveness of pollution prevention and control technologies, practices, programs, and regulations, along with socio-economic considerations, and indicate where there is a local, regional or national need for enhanced protection.

The EEM program has been designed on the basis of the non-duplication principle and to minimize overlap with existing provincial or local monitoring programs already in place. Monitoring data collected in a program required by another regulatory agency are acceptable if they are meeting the objectives of the EEM program and answer the key questions of data analysis and interpretation.

The National EEM Office within Environment Canada co-ordinates the assessment of the results of the EEM program on a national basis and the management of the data archive for the EEM program. In addition, the National EEM Office will form the EEM Co-ordination Group to review the overall effectiveness of the program, approximately three years after the commencement of the EEM program.

1.8.5 Provincial/Territorial Governments

The provinces/territories will be invited to provide a member to participate on the TAP and will work with the federal government to minimize duplication of programs.

1.9 References

Aquatic Effects Technology Evaluation (AETE). 1999. Report of Selected Technologies for Cost-effective Environmental Monitoring of Mine Effluent Impact in Canada. AETE synthesis Report. Prepared by ESG International Inc. for Aquatic Effects Technology Evaluation (AETE) Program, CANMET, Natural Resources Canada. Draft, February 1999.

AQUAMIN. 1996. Assessment of Aquatic Effects of Mining in Canada. Environment Canada.

Table of Contents

2. Study Design *2-1*

2.1 Overview *2-1*
2.1.1 First EEM Study 2-2
2.1.2 Second and Subsequent EEM Study Designs 2-3
2.1.3 Final Biological Monitoring Study Design (prior to a mine closing) *2-4*

2.2 Summary of Information to be Reported in Study Design for Biological Monitoring *2-4*
2.2.1 Site Characterization 2-4
2.2.2 Fish Survey 2-5
2.2.3 Fish Tissue Analysis 2-6
2.2.4 Benthic Invertebrate Community Survey 2-6
2.2.5 Alternative Methods for Biological Monitoring 2-7

2. Study Design

2.1 Overview

The objective of a study design is to describe how the biological monitoring studies (a fish survey - Chapter 4, fish tissue analysis - Chapter 4 or a benthic invertebrate community survey - Chapter 5) will be conducted, including the relevant site characterization information (Chapter 3), the dates and times that the information will be collected, quality assurance and quality control measures, and a summary of any previous biological monitoring information.

Most EEM studies involve collecting samples in both the exposure and reference areas, or along an exposure gradient (in the case of a benthic invertebrate community). When conducting an Investigation of Cause study, the mine may not necessarily be sampling an exposure and reference area. The **exposure area** is all fish habitat and waters frequented by fish that are exposed to mine effluent. Techniques for delineation of exposure areas are described in Section 3.3. For Initial and Periodic Monitoring studies, sampling in the exposure area is done in an area proximate to the effluent discharge, with the appropriate habitat for the benthic invertebrate community or where the selected fish species are present. In Focussed Monitoring studies, additional sampling within the exposure area will be conducted to determine the magnitude of the effects and determine the geographic extent. In order to determine the geographic extent of the effect, sampling is done progressively further from the effluent discharge until no effect is found.

The **reference area** is defined as waters frequented by fish that is not exposed to effluent and that has fish habitat that, as far as practicable, is most similar to that of the exposure area. The reference area should be located in the same water body as the effluent discharge, but upstream of or beyond any influence from the discharge. When a mine is located at a headwater, and/or where no suitable reference area on the same water body is available, then the reference area should be located in an adjacent water body with similar characteristics or an unimpacted tributary to the receiving water body. More than one reference area may be used, where appropriate. During Focussed Monitoring, it may be necessary to sample more than one reference area if different habitat types in the exposure area are sampled. As well, a more regional approach could be considered, particularly for benthic invertebrate community surveys, such as looking at several unimpacted streams in the area (Reference Condition Approach).

Even when pre-mining data are available, field sampling should be done at exposure and reference areas, and should be done at the same time, in order to ensure that reference conditions have not changed and that changes observed are not incorrectly attributed to the mine effluent.

Where possible, sampling areas for different components (fish, benthic invertebrate community, water) should be co-located. In addition, mines are encouraged to conduct benthic invertebrate community and fish monitoring studies concurrently. Where studies are proposed jointly by more than one mine, sampling areas may be shared.

Where historical monitoring data exist, the mine should consider using the same sampling areas providing they are appropriate for use in the EEM program. This will help ensure that monitoring data collected as part of the EEM program may be compared with historical data.

2.1.1 First EEM Study

The first EEM study design will be submitted to the Authorization Officer either:

a) 12 months after the mine becomes subject to Section 7 of the MMER; or
b) 24 months after the mine becomes subject to Section 7 of the MMER if the following conditions exist:
    i) the mine has conducted biological monitoring studies prior to the mine being subject to the MMER,
    ii) the previous biological monitoring studies have determined if the effluent is causing effects in fish, fish tissue or the benthic invertebrate community, and
    iii) the results of the biological monitoring studies were submitted in a report to the Authorization Officer within 12 months of the mine being subject to Section 7 of the MMER, that contained the scientific data to support the results.

The first EEM study design will outline the following information:

a) site characterization information (Chapter 3),

b) a description of how the fish survey (Chapter 4) will be conducted (if the concentration of effluent in the exposure area is greater than 1% in the area located within 250m of any final discharge point), that includes:

    - the fish species selected (Section 4.3.2.2.1);
    - the sampling areas selected (Section 4.2.2);
    - the sample size selected (Section 4.3.2.2.3);
    - the field and laboratory methodology selected (Section 4.6 - 4.9); and,
    - how the above will be used to provide the information necessary to determine if the effluent has an effect on the fish.

c) a description of how the fish tissue analysis will be conducted (if during effluent characterization (Chapter 6) the concentration of total mercury in the effluent is equal to or greater than 0.10 μg/L), that includes:

- the fish species selected (Section 4.3.2.2.1);
- the sampling areas selected (Section 4.2.2);
- the sample size selected (Section 4.10.1.1);
- the field and laboratory methodology selected (Section 4.6 & 4.10) ; and,
- how the above will be used to provide the information necessary to determine if the effluent has an effect on the fish tissue.

d) a description of how the benthic invertebrate community survey will be conducted, that includes:

- the sampling areas selected (Section 5.6 & 5.8);
- the sample size selected (Section 5.11);
- the sampling season selected (Section 5.10); and,
- the field and laboratory methodologies selected (Section 5.12, 5.21, 5.22); and
- how the above will be used to provide the information necessary to determine if the effluent has an effect on the benthic invertebrate community.

e) the dates and times that the samples will be collected for biological monitoring;

f) a description of the quality assurance and quality control measures that will be implemented to ensure the validity of the data collected; and,

g) a summary of the results of any previous biological monitoring studies.

2.1.2 Second and Subsequent EEM Study Designs

The second and any subsequent EEM study designs will be submitted to the Authorization Officer at least six months before the biological monitoring study is to be conducted.

For Periodic Monitoring, the second and subsequent study designs will include the same information that is included in the first EEM study design, including a summary of results of any previous EEM studies.

When the results of the two previous biological monitoring studies indicate an effect in fish, fish tissue or the benthic invertebrate community, a Focussed Monitoring study will be conducted. The study design for Focussed Monitoring will include the same information identified above for periodic monitoring study designs, and will also provide a description of one or more additional sampling areas within the exposure area, that will be used to assess the magnitude and geographic extent of the effect.

If the magnitude and geographic extent of the effects have been determined, but the cause of the effect is still unknown, the mine will conduct an Investigation of Cause study. The study design for Investigation of Cause may not necessarily include the same information that was included in the Periodic or Focussed Monitoring study designs. It will outline, in detail, what field and laboratory work will be conducted in order to determine the cause of the effect in fish, fish tissue or benthic invertebrate community.

2.1.3 Final Biological Monitoring Study Design (prior to a mine closing)

If a mine has submitted to the Authorization Officer the notice to close the mine under subsection 31(1) of the MMER, the mine will submit a study design within six months of providing the notice.

The study design will include the same information as was in the second or subsequent study design, depending upon which monitoring phase the mine is in (Periodic Monitoring, Focussed Monitoring, or Investigation of Cause).

2.2 Summary of Information to be Reported in Study Design for Biological Monitoring

2.2.1 Site Characterization

Detailed information on the site characterization information can be found in Chapter 3. The site characterization information that will be reported in the study design includes:

- a description of the manner in which the effluent mixes within the exposure area, including an estimate of effluent concentration at 250m from each planned final discharge point;
- a description of the geographical area where the EEM study will be conducted, including the reference and exposure areas, including applicable information on geological, hydrological, oceanographical, limnological, chemical and biological features;
- the type of production process used by the mine and environmental protection practices in place at the mine;
- a summary of any federal, provincial or other laws applicable to the mine in respect of effluent and environmental monitoring;
- a description of any factors, anthropogenic or natural, that are not related to the effluent under study, and may reasonably be expected to contribute to any observed effect; and,
- any additional information relevant to the completion of the site characterization.

2.2.2 Fish Survey

If an EEM study will include a fish survey (Chapter 4), then the study design will include a description and scientific rationale for the following:

The fish species selected (Section 4.3.2.2.1). It is recommended that two sexually mature fish species be selected. In selecting these two species, species selected for previous EEM studies should be considered, and preference should be given to:

- resident (non-migratory) fish species identified in the site characterization;
- sexually mature female and male fish species that are abundant in both the exposure and reference areas;
- fish species for which fishing or sampling permits can be obtained; and
- fish species that are most exposed to effluent in the exposure area.

If two sexually mature fish species cannot be selected for the fish survey, then one sexually mature fish species and one sexually immature fish species should be selected, or:

- two sexually immature fish species;
- one sexually mature fish species; or
- one sexually immature fish species.

The locations of sampling areas in the exposure and reference areas where the fish survey will be conducted. To select the location of the sampling areas (Section 4.2.2), the following factors should be considered:

- the area where fish are present in the exposure area that is most exposed effluent (Initial and Periodic Monitoring) or areas progressively further from the effluent discharge, in order to determine the geographic extent (Focussed Monitoring);
- the fish survey sample size;
- the abundance of the selected fish species in the potential sampling areas;
- the location of sampling areas selected in a previous EEM study; and
- the location of confounding influences.

The selected fish sample size (Section 4.3.2.2.3). The following sample sizes are recommended:

- 20 sexually mature males and 20 sexually mature females in each sampling area, and
- an additional 20 sexually immature fish in each sampling area, if the maximum size of the fish species selected is approximately 150millimetres.

However, the mine may select an alternate sample size based on a statistical power analysis of available data. If an alternate sample size is selected, the values of alpha and beta should be set equal, and the effect size used in the determination of the sample size should be provided.

If non-lethal sampling is to be conducted, no less than 100 fish of each fish species, representing different size classes, should be collected in each sampling area.

The type of field sampling equipment or method (Sections 4.6-4.9) that will be used in the sampling areas and the date and approximate time when the selected fish species will be collected.

2.2.3 Fish Tissue Analysis

If mercury has been detected in the effluent at a concentration greater than or equal to 0.10 μg/L, the mine will conduct a fish tissue analysis. The study design will describe and provide a scientific rationale for the following:

The fish species selected (Section 4.3.2.2.1). It is recommended that the fish species for the metal analysis be chosen based on what is locally consumed. As well, the tissue used for the analysis should be a tissue that is locally consumed.

The locations of sampling areas in the exposure and reference areas where the fish survey will be conducted. In selecting the location of the sampling areas (Section 4.2.2), the same factors considered for the fish survey should be taken into consideration.

The selected fish sample size (4.10.1.1). It is recommended that a composite sample of eight fish of the same sex and species be used.

The type of field sampling equipment or method (Sections 4.6 & 4.10) that will be used in the sampling areas; and the date and approximate time when the selected fish species will be collected.

2.2.4 Benthic Invertebrate Community Survey

For the benthic invertebrate community survey, the study design will describe and provide a scientific rationale for the following:

The sampling areas selected (Section 5.6 & 5.8).. In selecting the sampling areas, preference should be given to the benthic invertebrate habitat that:

- is identified in the site characterization and is predominant in the study area;
- has the greatest diversity of benthic invertebrate communities in the study

area; and
- is most exposed to the effluent in the exposure area (for Initial and Periodic Monitoring) or areas progressively further from the effluent discharge, in order to determine the geographic extent (Focussed Monitoring).

In selecting the sampling areas, the mine should take into consideration the sampling design that will be used, which may include: control/impact; multiple-control/impact; simple gradient; radial or multiple gradient; and reference condition approach.

In addition, the following factors should be considered:

- benthic invertebrate community sampling areas selected in a previous EEM study;
- the location of confounding influences; and
- the sampling design that will be used for the benthic invertebrate community survey.

The sample size selected (Section 5.11). Selection of the sample size (e.g., number of sampling stations) should be based on a statistical power analysis.

The values of alpha and beta should be set equal, and the effect size used in the determination of the sample size should be provided.

It is recommended that the study design also identify the number of benthic invertebrate community samples that will be collected at each sampling station and provide the rationale for selecting the number of samples that will be collected.

The sampling season for the benthic invertebrate community survey (Section 5.10). In selecting the sampling season, the sampling season selected in a previous EEM study should be considered, and preference should be given to:

- the season when the benthic invertebrate community is most diverse; and
- the season when the benthic invertebrate community's exposure to the effluent is expected to be at its highest.

Field program and laboratory work elements (Section 5.12, 5.21, 5.22):

- the field sampling equipment that will be used to collect samples at each of the sampling stations;
- the field sampling method that will be used at each of the sampling stations;
- the laboratory method that will be used to prepare samples for the purposes of identifying and enumerating the benthic invertebrates sampled; and,
- the method of sample collection and associated laboratory work for sediment samples to be collected at each benthic invertebrate community sampling station.

2.2.5 Alternative Methods for Biological Monitoring

There may be situations where the fish survey, using wild finfish as outlined above, may provide minimal useful information or may not be cost effective. This includes situations where:

- conducting the fish survey and fish tissue analysis in the study area would subject a person to hazardous conditions; or
- the fish survey and the fish tissue analysis cannot be used to evaluate an effect of effluent on the fish that is scientifically relevant as a result of confounding influences in the study area.

Chapter 12 outlines in detail the recommended methodologies for the alternative study designs. If an alternative study is to be conducted for the fish survey or fish tissue analysis, then the study design will describe in detail

- the fish species selected;
- the sampling areas selected;
- the sample size selected;
- the field and laboratory methodology selected; and,
- how the above will be used to provide the information necessary to determine if the effluent has an effect on the fish and/or the fish tissue.

If an alternative study to the benthic invertebrate community survey is to be conducted then the study design shall describe in detail:

- the sampling areas selected;
- the sample size selected;
- the sampling season selected;
- the field and laboratory methodologies selected; and.
- how the information will be used to provide the information necessary to determine if the effluent has an effect on the benthic invertebrate community.

Table of contents

3. Site Characterization *3-1*

3.1 Overview *3-1*

3.2 Habitat Mapping and Classification *3-5*
3.2.1 Environment-Specific Considerations 3-6
3.2.2 Delineation of Depositional Zones 3-9
3.2.3 Aquatic Resource Inventory 3-9

3.3 Effluent Mixing in the Exposure Area *3-12*
3.3.1 Measuring of Constituents in the Effluent 3-13
3.3.1.1 Conductivity Surveys 3-13
3.3.1.2 Tracing By Effluent Metals 3-13
3.3.2 Effluent Plume Delineation Studies 3-14
3.3.2.1 Using Tracers 3-15
3.3.2.2 Measurements in the Field 3-16
3.3.2.3 Environment-Specific Considerations 3-18
3.3.3 Using Predictive Models 3-22

3.4 Use of Sublethal Toxicity for Site Characterization *3-23*

3.5 Characteristics of Mining Environments *3-23*
3.5.1 Headwater Location of Mine Sites 3-23
3.5.2 Size and Life-Spans of Mines 3-24
3.5.3 Effluent Quality and Volume 3-24
3.5.4 Local Geology 3-25
3.5.5 Bioavailability of Metals 3-25

3.6 References *3-25*

List of Tables

Table 3-1: Site Characterization information for the preparation of an EEM study design 3-4
Table 3-2: Substrate subclasses and dominance types (after Cowardin et al. 1979) 3-8
Table 3-3: Salinity modifiers (after Cowardin et al. 1979) 3-8
Table 3-4: Water regime modifiers (after Cowardin et al. 1979) 3-8
Table 3-5a: Endangered* fish species in Canada (November 1998) 3-9
Table 3-5b: Threatened* fish species in Canada (November 1998) 3-10
Table 3-5c: Vulnerable* fish species in Canada (November 1998) 3-11

List of Figure

Figure 3-1: Classification hierarchy of wetlands and deep water habitats showing systems, subsystems, and classes (after Cowardin et al. 1979). 3-7

3. Site Characterization

3.1 Overview

Site characterization information will be submitted as part of each EEM study design. As identified in *Section 11 of Schedule 5* of the MMER. The information for site characterization will be included in detail in the first EEM study design. Subsequent EEM study designs may include the information in a summary format, updating any new information in detail.

In most cases, mines will have most site characterization information available from previous assessments and historical studies. Some additional field data may be required to provide adequate background for the first EEM study design, particularly with respect to hydrology and aquatic resources. Such information is critical to the design of EEM studies.

Site characterization information is used to identify suitable sampling areas with similar habitats in the exposure and reference areas, and to obtain information on other discharges and confounding factors that may affect the interpretation of data obtained from those areas. Information on some of the unique environmental characteristics of mine sites that should be taken into consideration during the site characterization, can be found in Section 3.5.

Much of the information for site characterization can be effectively reported in map form. Maps should be of sufficient scale (e.g. 1:5000) to show the features of the study area in adequate detail. The actual scale should be reported on any map used. The geographic extent of the study area to be mapped should be determined on a site-specific basis, and should include exposure and reference areas.

The site characterization section of the EEM study design will include the following information:

1) A description of the manner in which the effluent mixes within the exposure area (Section 3.3), including an estimate of the concentration of effluent in water 250 m from each final discharge point (as defined by the MMER).

2) A description of the reference and exposure areas where the biological monitoring studies will be conducted that includes information on the geological, hydrological, oceanographical, limnological, chemical and biological features of those areas (Table 3.1).

3) A description of the type of production process used by the mine and the environmental protection practices in place at the mine (Table 3.1);

4) A summary of any federal, provincial or other laws applicable to the mine in respect of effluent and environmental monitoring;

5) A description of any anthropogenic, natural or other factors that are not related to the effluent under study and that may reasonably be expected to contribute to any observed effect (Table 3.1); and,

Any other additional information relevant to the site characterization, which may include:

a) An identification of the major chemical reagents used in the overall production process since January 1, 1996. Mines are encouraged to report current quantities of reagents used. This should list reagents of the following types:
   - activators
   - flocculants
   - pH modifiers
   - depressants
   - frothers
   - collectors

b) An inventory of all discharge points where effluent is deposited into the exposure area. This inventory should identify all known sources of effluent to the aquatic environment, including those regulated under the MMER, as well as any others (e.g., non-point sources) which may have the potential to cause an effect on the aquatic environment.

c) Information on the local climate, particularly seasonal precipitation patterns.

Table 3-1: Site Characterization information for the preparation of an EEM study design

Information Type	**Recommended information to be reported**
General Characteristics	• bedrock and surficial geology • topography • soil and vegetation • site accessibility • climatology*
Hydrology	• watershed(s) description • water flow (rivers) or dispersion (lakes, estuaries, marine) characteristics* • general description of how effluent(s) mix(es) with receiving water* • bathymetry mapping (including slope in marine environments) • gradient (rivers) • tides (marine) - mean monthly tide height data • stratification patterns (thermal and chemical) • natural barriers to fish movement
Anthropogenic Influences	• docks, wharves, ferry terminals, marinas, boat launches, and public recreational zones • bridges, crossings and fordings • water intakes, effluent discharges, stormwater discharges, sewer overflows • waste disposal sites • contaminant source inventory, including point and non-point sources • dams, culverts, waterfalls and other barriers to fish movement • surrounding land use • location of aquaculture facilities

Aquatic Resource Characteristics	• location of exposure and reference areas used in historical studies
	• fish and shellfish species present (resident and migratory)*
	• relative abundance of fish and shellfish species*
	• utilisation of the exposure and reference areas by fish and shellfish (spawning grounds, nursery areas, etc.)
	• rare, threatened or endangered fish species (if present)*
	• non-commercial fisheries (recreational and subsistence)*
	• commercial fisheries*
	• zones of macrophyte growth
	• ecologically relevant benthic invertebrate habitat(s) and their relative proportions including:
	  • delineation of depositional and erosional zones
	  • substrate classification

Environmental Protection Systems and Practices	• water management* • effluent treatment* • residence time* • management of tailings* • polishing ponds* • waster rock* (including use of waste rock for mine back fill and building material)

[1] This information can effectively be reported in map form except for the characteristics identified by an (*). Sufficient scale should be used to show features in adequate detail, with actual scale reported on any map used.

3.2 Habitat Mapping and Classification

Some elements of habitat mapping and classification, as well as aquatic resource inventory, are included as part of site characterization. More detailed habitat mapping may be helpful in identifying habitat types present in the exposure and reference areas. This section provides guidance on habitat mapping and classification.

The recommended method to create a habitat map is to perform a habitat classification. The recommended framework (Figure 3.1) for classifying aquatic features is the classification system developed by the U.S. Fish and Wildlife Service (FWS), "Classification of Wetlands and Deepwater Habitats of the United States" (Cowardin et al. 1979, Busch and Sly 1992). The system allows for classification of a wide range of continental, aquatic, and semi-aquatic habitats.

The FWS classification system is hierarchical. Habitats are divided among five major systems at the broadest level: marine, estuarine, riverine, lacustrine, and littoral (Figure 3.1). Systems are further subdivided into subsystems that reflect hydrologic conditions, e.g. littoral vs. limnetic in the lacustrine system. However, the littoral system is not divided into subsystems. Below the subsystem level is the class level, which describes the appearance of the habitat in terms of vegetation (e.g. emergent wetland, aquatic bed, forested wetlands) or in terms of substrate where vegetation is inconspicuous or absent (e.g. unconsolidated shore, rocky shore, streambed). Each class can be further subdivided into subclasses and subclasses divided into dominant types.

For the purposes of site characterization, it is recommended that the classification system be used down to the class level, as shown in Figure 3.1. In addition, the FWS Subclass and Dominance Types for substrate and modifiers should be used for salinity and water regimes (Tables 3.2 to 3.4). Because the FWS system provides little or no resolution in the vertical (depth) dimension, collection of information on temperature and oxygen profiles during limiting times of the year (generally summer or possibly under ice in winter), and both vertical and horizontal salinity gradients under various tidal conditions is recommended if the receiving system is subject to vertical stratification (e.g. estuarine).

Additional guidance on habitat description procedures for coastal and estuarine situations is provided in Cowardin et al. (1979). In addition, it is recommended that bottom substrates be described, where appropriate. Further guidance on aquatic habitat assessment can also be found in the Department of Fisheries and Oceans and the British Columbia Ministry of the Environment and Parks (1987), Orth (1989), Ontario Ministry of Natural Resources (1989), Plafkin et al. (1989), and the Department of Fisheries and Oceans (1990).

3.2.1 Environment-Specific Considerations

Rivers: It is recommended that river habitat descriptions include information on elevation gradient; the location of dams, falls and other barriers to fish migration; mean annual discharge and ranges; and general substrate characteristics of each river (preferably in the form of a gradient profile chart). Upstream and downstream inputs (e.g. storm water, sewer overflow, effluent from other industrial sites) should be mapped and described.

Lakes: Important habitat features of lakes include bathymetry, the locations of major inlets and outlets, and general oxygen-temperature conditions (e.g. thermal stratification, occurrences of oxygen depletion in deep water).

Open Coastlines: Suggested additional mapping parameters for open coastlines (marine, Great Lakes) include depth contours, near-shore substrate characteristics, shoreline configuration, and the locations of inflowing rivers and other discharges and activities.

Estuaries: Estuaries are best described in terms of their general salinity gradients, flows, bathymetries, and general substrate features. A description of tidal cycles is recommended for all marine and estuary locations. Most of the above features can be described from navigational maps, topographic maps, government publications on tides and river discharge records, and through interviews with local government officials and knowledgeable individuals.

Figure 3-1: Classification hierarchy of wetlands and deep water habitats showing systems, subsystems, and classes (after Cowardin et al. 1979).

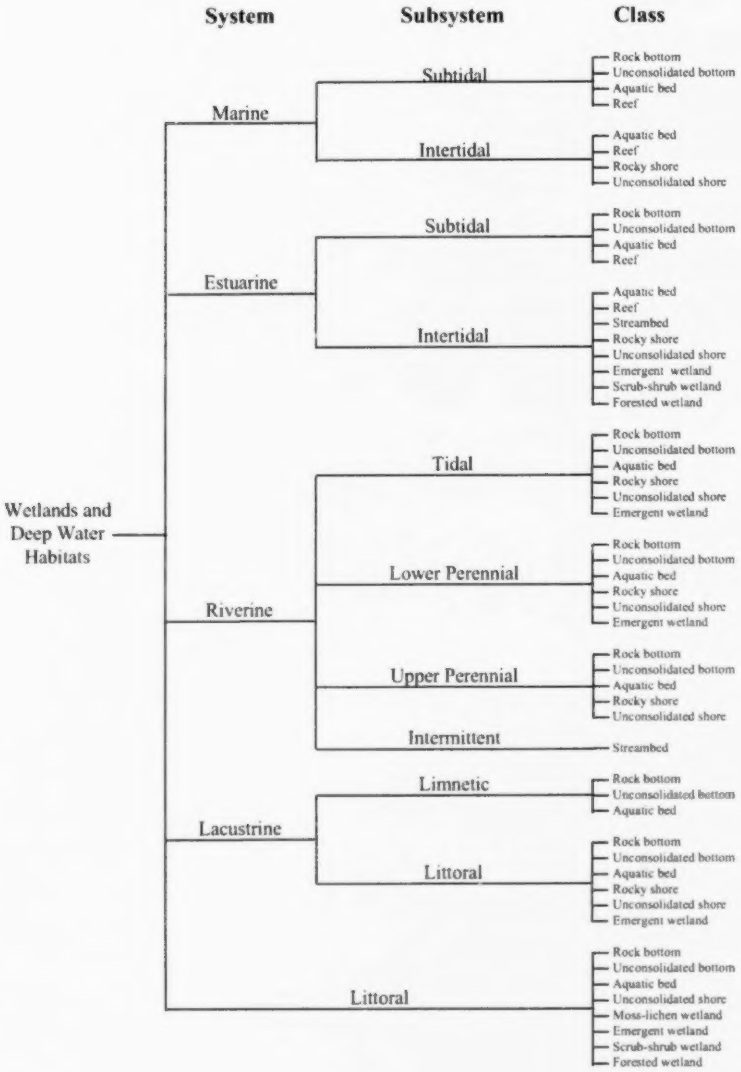

System
Subsystem
Class
Wetlands and Deep Water Habitats
Marine
Subtidal
Rock bottom
Unconsolidated bottom
Aquatic bed
Reef
Intertidal
Aquatic bed
Reef
Rocky shore
Unconsolidated shore
Estuarine
Subtidal
Rock bottom
Unconsolidated bottom
Aquatic bed
Reef
Intertidal
Aquatic bed
Reef
Streambed
Rocky shore
Unconsolidated shore
Emergent wetland
Scrub-shrub wetland
Forested wetland
Riverine
Tidal
Rock bottom
Unconsolidated bottom
Aquatic bed
Rocky shore
Unconsolidated shore
Emergent wetland
Lower Perennial
Rock bottom
Unconsolidated bottom
Aquatic bed
Rocky shore
Unconsolidated shore
Emergent wetland
Upper Perennial
Rock bottom
Unconsolidated bottom
Aquatic bed
Rocky shore
Unconsolidated shore
Intermittent
Streambed
Lacustrine
Limnetic
Rock bottom
Unconsolidated bottom
Aquatic bed
Littoral
Rock bottom
Unconsolidated bottom
Aquatic bed
Rocky shore
Unconsolidated shore
Emergent wetland
Littoral
Rock bottom
Unconsolidated bottom
Aquatic bed
Unconsolidated shore
Moss-lichen wetland
Emergent wetland
Scrub-shrub wetland
Forested wetland

Table 3-2: Substrate subclasses and dominance types (after Cowardin et al. 1979)

Substrate Class	Subclass/Dominance Type	Substrate Class	Subclass/Dominance Type
Rock bottom	bedrock rubble	Unconsolidated bottom	cobble-gravel sand mud organic
Aquatic bed	Algal Aquatic moss rooted vascular plants floating vascular plants	Reef	Mollusc
Streambed	bedrock rubble cobble-gravel sand mud organic	Rocky shore	Bedrock rubble

Table 3-3: Salinity modifiers (after Cowardin et al. 1979)

Modifiers	Salinity (parts per thousand)	Approximate Specific Conductance (μmhos at 25°C)
Hyperhaline	>40	>60,000
Euhaline	30-40	45,000-60,000
Mixohaline (Brackish)	0.5-30	800-45,000
Polyhaline	18.0-30	30,000-45,000
Mesohaline	5.0-18	8,000-30,000
Oligohaline	0.5-5	800-8,000
Fresh	<0.5	<800

Table 3-4: Water regime modifiers (after Cowardin et al. 1979)

Subclass	Modifier	Subclass	Modifier
Tidal	Subtidal Irregularly exposed Regularly flooded Irregularly flooded	Nontidal	Permanently flooded Intermittently exposed Semi-permanently flooded Seasonally flooded Temporarily flooded Intermittently flooded

3.2.2 Delineation of Depositional Zones

Depositional zones occur where water velocity decreases resulting in particles settling out; the finest particles settle out in the slowest current speeds. Historical contaminant or benthic invertebrate community data may be helpful in identifying sampling stations within the exposure area (Section 3.3). To compare resident benthic invertebrate communities, similar (but uncontaminated) sediment depositional zones should be located in the reference area.

Depositional zones in the exposure area will be identified and illustrated on the habitat map. Any information on sediment characterization (chemistry, toxicity) should be reported.

3.2.3 Aquatic Resource Inventory

An aquatic resources inventory includes the identification of fish and shellfish (resident and transient) that are presently being fished commercially and non-commercially (sport, including stocked fish, and subsistence fishery). In addition, any species recognized by federal, provincial or territorial authorities as rare, threatened or endangered (Table 3.5) as well as any species which may be present in sufficient numbers to be considered as a fish species appropriate for EEM must be documented within the site characterization, as described in Section 3.1. This information can be obtained from district fisheries biologists in federal, provincial or territorial regulatory or museum agencies, from local conservation officials and from interviews with members of the local community (fishermen, aboriginal and public interest groups). Aquaculture installations should also be noted.

The success of subsequent field programs increases with familiarity of the study area. It is recommended that field work be undertaken to verify historical information if this information is not detailed or recent.

The study design should consider the utilization of the exposure area by fish species in order to suitably design the study. For example, an exposure area predominantly used by juvenile fish in which adults are transient will require an equivalent reference area and dictate the measurement of the appropriate indicators. In this case the program would be assessing differences in growth. Reproduction would not be measured due to the absence of gonads in juvenile fish. Stocked fish are not appropriate for EEM type monitoring, as these fish are predominately sport fish. As well, stocked fish generally have no apparent reproductive success, therefore this key endpoint could not be evaluated.

Table 3-5a: Endangered* fish species in Canada (November 1998)

Provinces/Territories	Species Names		Last Examined
British Columbia	Salish Sucker	*Catostomus* sp.	1986
	Nooksack Dace	*Rhinicthys* sp.	1996
Nova Scotia	Acadian Whitefish	*Coregonus huntsmani*	1987
Ontario	Aurora Trout	*Salvelinus fontinalis timagamiensis*	1984

* Endangered species: a species facing imminent extirpation or extinction (COSEWIC, 1998).

Table 3-5b: Threatened* fish species in Canada (November 1998)

Provinces/Territories	Species Names		Last Examined
Alberta	Shortjaw Cisco	*Coregonus zenithicus*	1987
	Great Lakes Deepwater Sculpin	*Myoxocephalus thompsoni*	1987
British Columbia	Shorthead Sculpin	*Cottus confusus*	1984
	Texada Stickleback - benthic and limnetic form	*Gasterosteus* sp.	1998
	Enos Lake Stickleback 2 species	*Gasterosteus* spp.	1988
Manitoba	Shortjaw Cisco	*Coregonus zenithicus*	1987
	Bigmouth Buffalo	*Ictiobus cyprinellus*	1989
	Great Lakes Deepwater Sculpin	*Myoxocephalus thompsoni*	1987
New Brunswick	Lake Utopia Dwarf Smelt	*Osmerus* sp.	1998
Northwest Territories	Shortjaw Cisco	*Coregonus zenithicus*	1987
	Great Lakes Deepwater Sculpin	*Myoxocephalus thompsoni*	1987
Ontario	Eastern Sand Darter	*Ammocrypta pellucida*	1994
	Lake Simcoe Whitefish	*Coregonus clupeaformis*	1987
	Blackfin Cisco	*Coregonus nigripinnis*	1988
	Shortnose Cisco	*Coregonus reighardi*	1987
	Shortjaw Cisco	*Coregonus zenithicus*	1987
	Bigmouth Buffalo	*Ictiobus cyprinellus*	1989
	Black Redhorse	*Moxostoma duquesenei*	1988
	Great Lakes Deepwater Sculpin	*Myoxocephalus thompsoni*	1987
	Margined Madtom	*Noturus insignis*	1989
	Channel Darter	*Percina copelandi*	1993
Quebec	Eastern Sand Darter	*Ammocrypta pellucida*	1994
	Copper Redhorse	*Moxostoma hubbsi*	1987
	Great Lakes Deepwater Sculpin	*Myoxocephalus thompsoni*	1987
	Margined Madtom	*Noturus insignis*	1989
	Channel Darter	*Percina copelandi*	1993
Saskatchewan	Shortjaw Cisco	*Coregonus zenithicus*	1987
	Bigmouth Buffalo	*Ictiobus cyprinellus*	1989
	Great Lakes Deepwater Sculpin	*Myoxocephalus thompsoni*	1987

* Threatened species: a species likely to become endangered if limiting factors are not reversed (COSEWIC, 1998).

Table 3-5c: Vulnerable* fish species in Canada (November 1998)

Provinces/Territories	Species Names		Last Examined
Oceans			
Arctic	Blackline Prickleback	*Acantholumpenus mackayi*	1989
	Bering Wolffish	*Anarhichas orientalis*	1989
Atlantic	Atlantic Cod	*Gadus morhua*	1998
Pacific	Pacific Sardine	*Sardinops sagax*	1987
Alberta	Western Silvery Minnow	*Hybognathus argyritis*	1997
British Columbia	Green Sturgeon	*Acipenser medirostris*	1987
	White Sturgeon	*Acipenser transmontanus*	1990
	Cultus Pygmy Sculpin	*Cottus* sp.	1997
	Charlotte Unarmoured Stickleback	*Gasterosteus aculeatus*	1983
	Giant Stickleback	*Gasterosteus* sp.	1980
	Lake Lamprey	*Lampetra macrostoma*	1998
	Speckled Dace	*Rhinichthys osculus*	1980
	Umatilla Dace	*Rhinichthys umatilla*	1988
Manitoba	Chestnut Lamprey	*Ichthyomyzon castaneus*	1991
	Northern Brook Lamprey	*Ichthyomyzon fossor*	1991
	Silver Chub	*Macrhybopsis storeriana*	1985
	Bigmouth Shiner	*Notropis dorsalis*	1985
	Rosyface Shiner	*Notropis rubellus*	1994
New Brunswick	Shortnose Sturgeon	*Acipenser brevirostrum*	1980
	Redbreast Sunfish	*Lepomis auritus*	1989
Newfoundland	Banded Killifish	*Fundulus diaphanus*	1989
Northwest Territories	Fourhorn Sculpin	*Myoxocephalus quadricornis*	1989
Ontario	Redside Dace	*Clinostomus elongatus*	1987
	Kiyi	*Coregonus kiyi*	1988
	Lake Chubsucker	*Erimyzon sucetta*	1994
	Greenside Darter	*Etheostoma blenniodes*	1990
	Blackstripe Topminnow	*Fundulus notatus*	1985
	Northern Brook Lamprey	*Ichthyomyzon fossor*	1991
	Black Buffalo	*Ictiobus niger*	1989
	Spotted Gar	*Lepisosteus oculatus*	1994
	Warmouth	*Lepomis gulosus*	1994
	Orange Spotted Sunfish	*Lepomis humilis*	1989
	Silver Chub	*Macrhybopsis storeriana*	1985
	Spotted Sucker	*Minytrema melanops*	1994
	River Redhorse	*Moxostoma carinatum*	1987
	Pugnose Shiner	*Notropis anogenus*	1985
	Silver Shiner	*Notropis photogenis*	1987
	Brindled Madtom	*Noturus miurus*	1985
	Northern Madtom	*Noturus stigmosus*	1998
	Pugnose Minnow	*Opsopoedus emilae*	1985
Quebec	Spring Cisco	*Coregonus* sp.	1992
	Northern Brook Lamprey	*Ichthyomyzon fossor*	1991
	River Redhorse	*Moxostoma carinatum*	1987
Saskatchewan	Western Silvery Minnow	*Hybognathus argyritis*	1997
	Chestnut Lamprey	*Ichthyomyzon castaneus*	1991
Yukon Territory	Squanga Whitefish	*Coregonus* sp.	1987

* Vulnerable species: a species of special concern because of characteristics that make it particularly sensitive to human activities or natural events (COSEWIC, 1998)

3.3 Effluent Mixing in the Exposure Area

During site characterization, the mine will describe in the study design, the manner in which effluent mixes within the exposure area. This description should include an indication of relative flow of the effluent and receiver, as well as seasonal variations in flow. This will give an indication of dilution rate. The description should also give an indication of the density of the effluent, and where within the water column the effluent is likely to be, prior to complete mixing. This estimate may be based on direct measurements in the field or modelling. Details on possible field measurements are provided in this section.

If the effluent concentration is less than 1% at 250m from each final point of effluent discharge, the fish population survey and fish tissue analysis is not conducted. In this case, more rigorous plume delineation methods are recommended to document the effluent concentrations in the exposure area.

It is recommended that the description of the manner in which effluent mixes within the exposure area, which is required as part of site characterization, include the following:

- identification of where in the exposure area the effluent is located, prior to mixing with the receiving water;
- estimation of where in the exposure area the effluent and receiving water begin mixing, and where mixing is complete;
- estimation of the effluent dilution ratio at points downstream of effluent discharge; and,
- identification of significant sources of dilution, other than the primary receiver (i.e. tributaries or other streams).

The following sections describe methods which can be used to obtain information needed to describe effluent mixing. There is also extensive guidance on plume delineation, in the event that more detailed information is necessary. Guidance is also provided on the use of predictive models, although, if modelling is used, it is recommended that some field measurements be taken to verify the model predictions.

3.3.1 Measuring of Constituents in the Effluent

3.3.1.1 Conductivity Surveys

If the conductivity of effluent is consistent during the period of study, then a conductivity survey can be used to locate the effluent plume within the exposure area. Survey results can be assessed semi-quantitatively, or conductivity measurements can be converted into relative effluent concentrations lying between 1 (effluent) and zero (background) by applying the following formula:

$$C_r = (C_a \times C_B)/(C_e \times C_b)$$

Where: C_e = effluent conductivity, µS/cm
C_b = background conductivity, µS/cm
C_a = measured conductivity to be converted, µS/cm
C_r = relative concentration

The relative concentration is an expression of the dilution ratio. Temperature readings need to be taken along with the measurements obtained from the conductivity meter, since conductivity rises approximately 2% for every $1^{\circ}C$ rise in temperature. Further information on theoretical considerations for effluent plume delineation using conductivity can be found in relevant reference material (e.g., Fischer et al. 1979; Freeze and Cherry 1979).

Although conductivity surveys can provide valid, cost effective estimates of the location of effluent within the receiving environment, the natural variability of conductivity in surface waters can interfere with locating the edge of the effluent plume. Such natural variability in conductivity may be observed with depth as well as for surface measurements. The presence of multiple tributaries or receiving water bodies can further exacerbate this difficulty.

3.3.1.2 Tracing By Effluent Metals

The location of the effluent plume within the exposure area may also be estimated by choosing a reference parameter present in the effluent and tracking its fate in time and space by measuring its concentrations in water samples taken at specific locations. The selection of such a tracer needs to be based on its stability and consistency of concentration, as well as on representativeness and ease of measurement. Since the parameter selected has to be a conservative substance, metals such as copper or nickel may serve as a "natural" tracer. Sulphate is often a good tracer of base metal mine effluents, particularly in massive sulphide deposits.

It should be cautioned that the effluent parameter selected for tracing may sometimes be present at the same order of magnitude in receiving waters. This would rule out its application for determining effluent dilution. Other parameters may be specific to the effluent alone, but present in such low concentrations as to make their detection difficult. Several parameters may be present at significantly greater levels in effluent, which would make them ideal tracers for dilution measurement. However, as a result of such complications as the costs of analysis, instability of the substance, difficulty in measurement, or the lack of an adequate *in situ* measuring apparatus, these parameters may not always prove to be appropriate "natural" tracers. The potential for effluent metals to be used as tracers for plume delineation should therefore be evaluated on a case- specific basis.

3.3.2 Effluent Plume Delineation Studies

Effluent plume delineation studies can be used to quantitatively assess effluent dispersion and dilution within the exposure area. Plume delineation studies are usually conducted by measuring effluent concentrations of tracers. Each type of receiving environment will have different types of characteristics to be considered in the design of these studies. General principles for conducting an effluent plume study are detailed below.

The recommended method to verify effluent dilution predictions is to measure at the middle of the initial discharge zone by sampling at the depth of highest concentration. By collecting three or four samples, the initial dilution can be calculated as the maximum measured tracer concentration divided by the tracer concentration in the effluent prior to discharge.

In some special cases, it may not be possible to measure the plume to an appropriate level of resolution of effluent concentration. The extent of the exposure area should be limited or expanded as appropriate, based on best professional judgement. For example, in a river with low flow, or in an exposure area where there are other discharges close to the mine effluent discharge, these factors will act to complicate concentration measurements and confound the results.

In some cases, it may not be possible or practical to use the effluent plume delineation methods described in this section. Examples of such situations include:

- Sites where the position of the effluent plume is highly transitory (i.e. the plume rapidly changes location, volume, or direction of flow)
- Sites where there is high dilution (i.e. where the effluent is diluted very rapidly near the discharge point and the locating of sampling stations in the mixing zone would be very difficult). In this case, it might be wise, where the effluent seems to dilute rapidly, to radiate out from the discharge point and keep looking for evidence of the plume beyond the very low concentrations. Likewise, when the dispersion of the plume is transitory, its maximal dispersion should be delineated, i.e. to include all areas of potential exposure.

3.3.2.1 Using Tracers

Type of Tracers

The ideal effluent plume tracer has a near zero background level and decay rate (a conservative substance), can be readily measured in the field at low concentrations, and can be released at a constant loading rate. These conditions are desirable because the extent of the effluent plume at the time of measurement should reflect near steady-state conditions during the period of measurement, and detection at low levels is needed to achieve the degree of resolution.

- ***Added Tracers:*** The preferred tracer is a fluorescent dye. The recommended dye of choice is Rhodamine WT. This dye has been shown not to be a mutagen (Douglas et al. 1983), has low potential for toxicity and adverse effects in the aquatic environment (Parker 1973), is safe when handled with care, generally available, can be readily measured in the field at concentrations in the low μg/L range, can be considered conservative in most cases, and typically has a near zero background level. Other tracers demonstrated as scientifically valid can also be used.

- ***Constituents in discharge:*** The effluents from some mines may contain metals that could potentially be used as tracers for delineating the zone of effluent mixing. However, some of these constituents can be highly variable in terms of their mass loading to the exposure area and are not easily measured in the field.

Mixing of Added Tracers

If tracers are to be added to effluent, they need to be added before the effluent is discharged. To ensure complete mixing of the tracer with the effluent before the effluent reaches the exposure area, the injection location should be well upstream from the effluent discharge point. If the dye is injected into an existing outfall, the turbulence in the outfall downstream from the injection point is sufficient to ensure complete mixing of dye throughout the effluent. The dilution of dye in the outfall is such that any change in effluent density caused by the addition of dye is insignificant. Small differences in temperature (2°C to 3°C) between the effluent and the dye can cause a significant density difference. If the tracer is not the same density as the effluent, it may separate from the effluent in the receiving water. The density of the dye Rhodamine WT should be adjusted with methyl hydrate or a similar solvent (dilution liquid) that has a specific gravity of less than 1.

Degradation of Tracers

Fluorescent tracers such as Rhodamine WT can be affected by oxidizing agents (e.g. bleaches, sulphides, sunlight, salt water, and microorganisms). Preliminary tests of dye-effluent interaction to describe the stability of the tracer and to determine the loss coefficient that should be used with the tracer are recommended.

The degradative effects of sunlight and microorganisms on the tracer can be determined by measuring the tracer concentration dynamics in test samples exposed to sunlight and by maintaining the effluent tracer mixture and receiving water at appropriate dilutions during the tracer studies. If the degradation of the dye is too large to be useful, another more conservative tracer may be required. Sulphur hexafluoride has been used and can be detected in very small concentrations using gas chromatography.

Injection System

A continuous flow constant rate injection system is preferred to simulate the operation of an outfall with continuous discharge loading. This type of injection system makes field measurements more reliable. The use of a constant rate positive displacement pump (e.g. peristaltic digital pump) is recommended. A constant head device is also a good method for injecting dye. The pumping rate should be checked at least at the beginning and the end of the injection by direct measurement. Effluent plume measurements can be repeated and mass balance checks can be carried out at cross plume transects in the plume. For a tracer with zero background levels (e.g. a dye), the injection rate can be calculated as follows:

$$q_{inj} = q_{eff} \times C_x \times 100 / [\%_{eff} \times C_o]$$

where:		
	q_{inj}	is the tracer injection rate, L/s
	q_{eff}	is the effluent flow rate, L/s
	C_x	is the tracer detection limit concentration, µg/L
	$\%_{eff}$	is the limit of the effluent plume in % effluent concentration
	C_o	is the concentration of the tracer in the injection mixture, µg/L

The concentration of the tracer should be considered in the computation of the injection rate. The concentration of the dye in the stock supplied by the manufacturer is defined. Any dilution of this stock should be considered in the computation of the injection rate to determine the concentration of the dye in the mixture injected.

Using the equation above, the constant rate injection of the tracer is determined by measuring the effluent discharge rate and then determining, as appropriate, the tracer injection rate based on the detection limits of the tracer measuring technique. For example, if the detection limit of the tracer measuring method is 1 µg/L, the limit of resolution of the effluent plume is 1% effluent concentration, the effluent flow is 1000 L/s and the tracer concentration in the injection mixture is 1 x 10^9 µg/L, then using the above equation the required pumping rate is 0.0001 L/s.

3.3.2.2 Measurements in the Field

The following section assumes the use of Rhodamine WT as the tracer. Use of an alternative tracer will require appropriate modifications to the recommended procedure. Details on the use of Rhodamine WT are available from the manufacturers of the fluorometers and the U.S.

Geological Survey (1986, 1989). For a mine with multiple discharges of process water, it is recommended that each discharge be dye-traced separately to define its individual effluent plume configuration.

The recommended method is as follows:

1) Carefully record the location of the tracer concentration measurements in the effluent plume. The precision of positioning systems would be determined on a site-specific basis.

2) Measure the tracer concentrations along lines perpendicular to the direction of travel of the effluent plume by traversing the plume in a boat equipped with a recording flow-through fluorometer and a pump sampling system, or preferably an *in situ* submersible fluorometer. A towed *in situ* fluorometer used in conjunction with an accurate real-time positioning system is recommended. Such a system can be repeatedly towed through the water column to rapidly build a three-dimensional picture of the effluent concentration.

   In situ fluorometers overcome many of the problems of a cheaper pumped sampling system, such as, mixing within the sampling hose, time delay caused by travel through the sampling hose, and the resulting slow response to concentration gradients. When the dye concentration profiles are measured *in situ* there is no time delay between the sample measurements and the location of the boat, compared to using a pumped sampling system. Grab samples are recommended only if the effluent plume is not accessible by boat. Grab samples are collected for subsequent analysis in the laboratory. At least 12 samples at each transect are recommended in the effluent plume to adequately define the plume configuration. This procedure is slower, has poor spatial resolution, and does not provide a continuous profile of the tracer concentrations. Consequently, it is difficult to carry out a mass balance of the tracer.

3) The depth of sampling should match the depth distribution of the tracer. The depth of maximum tracer concentration (trapping depth) is used for the concentration profiling of the tracer. By using an effluent tracer, the effects of stratification in the receiving water on the effluent are directly measured. Where the effluent is not well mixed with depth, the vertical profile should be defined. Towing should be done at the maximum concentration depth with periodic checks to ensure towing is at maximum concentration.

4) Sample the effluent plume at various distances downstream, approximately doubling the downstream distance intervals between transects to identify unusual characteristics in the plume. Unusual characteristics could include high concentrations observed at places beyond where the effluent concentration has dropped below the specified concentration (e.g. accumulation in an inlet, bay, etc.), or where an undertow moves effluent downstream or off-shore and resurfaces elsewhere, or similar situations.

5) Measure and record the bathymetry (depth sonar techniques are generally adequate) at tracer profiling locations. Determine positions of the transects using a locating system such as satellite fixing or optical fixing on shore landmarks. Where a detailed

hydrographic chart is already available, a survey of bathymetry may not be necessary. It is recommended that the bathymetry data of the receiving water be presented on a map of the exposure area.

6) If the effluent plume is influenced by other effluent plumes in the vicinity, design the water quality sampling program to consider these other plumes. For other discharges upstream, collect a water quality sample upstream from the target discharge and treat it as a background concentration for the effluent plume water quality.

3.3.2.3 Environment-Specific Considerations

For the purposes of this document, the type of exposure area environment is based on the most predominant characteristics as follows:

headwater:	current in one direction, effluent may contribute most of flow, no flow data
river:	current in one direction only, flow seasonally variable
lake:	currents in many directions
estuary:	stratified fresh water and salt water with tides
fjord:	relatively small fresh water discharges with tides
ocean:	sea water with tides and currents in many directions

Recommended methods for different exposure area types follow.

Headwaters

Water quality analysis and modelling often involve the consideration of the exposure area's minimum flow rate. A variety of minimum flow rates can be estimated from average daily, weekly, or monthly flow records. The frequency of occurrence of the flow must also be stated, that is, the minimum flow one might expect every 1, 2, 5, or 10 years. In general, for most water quality design work, the minimum average 7-day flow expected to occur once every 10 years is applied and is designated as the 7Q10 flow. The Ontario Ministry of Environment (MOE) requires application of the average minimum 7-day flow expected to occur once in 20 years (7Q20) for assimilative capacity studies. However, several different types of average minimum flows may be estimated for any given month, season, or time of year. Derivations of such flow rates generally assume the availability of long-term hydrologic records from established flow gauging stations. For lake systems, expected minimum lake elevations may need to be considered.

When hydrologic records for the exposure area are not available, which is commonly the case, other methods must be used to estimate minimum flow rates. These methods may involve correlation of the exposure area hydrology with that of a similar gauged system, or empirical flow estimation through knowledge of an areas meteorology, climate, and watershed characteristics. In North America, a considerable amount of research has been performed to derive relationships between rainfall, runoff, and stream flow. Such literature may be applied to estimate average minimum flow rates when little or no flow data are available for the water body in question. Difficulty arises when trying to determine the frequency of occurrence of flows estimated through empirical or rainfall-runoff relationships. Such estimates would have to be verified by a certain degree of ground truthing. Possible relationships to consider can be obtained by consulting available literature (Haan et al. 1982 and 1994; Viessman et al. 1989; Thomann and Mueller 1987).

Rivers

To bracket the range of expected conditions, effluent plume delineation studies should be carried out both during the annual low and maximum river flows. If one of these conditions occurs under an ice cover, the lowest or maximum annual river flow in the ice-free period should be used given operational and safety considerations. The continuous flow constant rate tracer injection technique should be used (see section 3.3.2.1). The number of measurements will be site-specific. It is recommended that concentrations be measured at various distances downstream to determine that the spatial extent of the effluent plume. Slug tests are also a suitable technique for monitoring mixing in rivers.

River flow for long-term average conditions can be interpolated from measured effluent plume configurations under the annual low and high river flow conditions (U.S. EPA 1987). Calculations and modelling can be used for this interpolation.

Lakes

The recommended method for the delineation of the effluent plume for a lake discharge is the constant tracer injection procedure. The slug test is also a suitable technique for monitoring mixing in lakes. General considerations for designing the study are as follows:

- the direction of movement and dilution characteristics of the effluent plume (variability depending on the prevailing currents);
- the effluent may not be well mixed with depth;
- most of the time the effluent plume for a mill discharge is near the water surface; and,
- the tracer has the same density as the effluent.

The tracer studies should be designed to fulfil the general requirements and to consider seasonal variations. The duration and timing of each study will be influenced by prevailing currents.

It is recommended that drogues be released and tracked concurrent with the dye tracing, to measure currents. The drogue water sails should be set 1.5 m below the water surface or set at the trapping depth of the effluent, which can be determined on site from the initial tracer monitoring. If clusters of drogues are released, dispersion for a batch release can be obtained from the paths of the individual drogues by determining the variance of the individual drogues about the centroid.

It is recommended that at least two recording current meters of appropriate sensitivity and a wind anemometer be operated for a period of a month for each season of measurement to monitor for seasonal variation in the vicinity of the discharge point. Data from the recording current meters are analyzed to quantify the statistical characteristics of the current climate in the vicinity of the discharge point. Site-specific evaluations of local airport or other weather data should be done to establish the usefulness of such data. The statistical relationship between the dilution and travel time is then combined with the statistical characteristics of the currents to develop spatial dilution envelopes (a graph showing the frequency of dilution factor as a function of two variables: dilution and travel time) about the discharge point.

The dilution envelopes should be developed and statistical confidence determined and recorded. These envelopes should show the probability of the effluent plume being present at any exposure area location and the mean and standard deviation of the effluent concentration at any location in the exposure area. The envelopes do not show the configuration of the effluent plume under specific current speed and direction conditions. To determine the effluent plume configuration at any particular time, and hence for a specific condition of current speed and direction, it is necessary to carry out a tracer study.

Unlike the river exposure area where the plume configuration can be predicted for the design condition using models, the design conditions for lakes are not easily predicted (see Section 2.2.3). Some provincial regulatory agencies specify conditions for which the effluent plume should be defined. For example, in Ontario, the Ministry of the Environment requires that the plume be defined for the high velocity condition normally greater than 20 cm/sec in both directions parallel to shore and the stagnation condition.

Estuaries

In most instances, industrial sites are located at the freshwater end of an estuary, primarily because the site requires process water from a freshwater source. There are exceptions where sites obtain their freshwater from other sources or small streams. In these instances, the ocean processes predominate and require special studies. These are discussed in the following as fjords.

In estuaries, in most cases, tidal effects and the depth distribution of the effluent, which is fresh water in salt water, are important characteristics. The currents in estuary receiving waters are to a large extent determined by the tidal variations.

Over the range of tidal conditions, tidal velocities are largest during the rising (flooding) and falling (ebbing) tides for the spring tidal condition. The effluent plume should be measured during maximum flood and ebb currents and at slack tide to delineate the minimum and maximum extent of the effluent plume. The rising and falling tide studies should be carried out during the period when river flow into the estuary is at an annual low.

The suggested method for estuary tracer studies is the continuous constant rate injection of a tracer for a period of about 4 hours starting at the beginning of the rising or falling tide. Mix the tracer with the effluent before discharge and measure the resulting tracer plume at right angles to the direction of movement of the plume and at various depths. To completely sample the plume, it may be necessary to carry out tracer injections over several tidal cycles. Salinity-temperature-depth (STD) data should also be collected in each of the tracer profile cross-sections. Records of tidal stage and currents in the areas affected by the plume should be kept during the tracer studies.

The measured plumes for the rising and falling tides represent the maximum upstream and downstream extent of the plume. These plumes are plotted as percentage effluent. The slack tide conditions are between the rising and falling tides and can be obtained by drawing envelopes about the data collected from this period.

Fjords

In a fjord, plume transport and dilution are caused primarily by residual currents. The residual current pattern can be determined in the vicinity of the discharge by injecting tracers into the effluent at a continuous rate over three tidal cycles, and measuring the resulting dye plume during slack tide after the second and third tidal cycles.

Concurrent with the tracer studies, at least two recording current meters operated at maximum concentration depth in the fjord are recommended during the study and for an additional 60-day period after completion of the study. Using the statistical data on the residual currents and measured tracer plume data, dilution envelopes can be developed.

Some fjords are very complex, and using locally measured plume characteristics and residual currents may not be sufficient to delineate the effluent plume, particularly for the far-field or for low effluent concentrations. In these cases it may be useful to combine a residual current transport model with a contaminant model such as the WASP (U.S. EPA 1988).

Normally, carrying out the studies during the spring rising and falling tides when the river flow is at its annual low flow is adequate for design conditions. Prediction models may be useful to determine the plume. Models for predicting the effluent plume for spring conditions are discussed in U.S. Army Corps of Engineers (1986) report on the CE-QUAL-W2 model.

Oceans

Ocean exposure areas are distinguished from estuaries in that there is no fresh water in the vicinity of the discharge. The effluent plume is diluted and transported by currents generated by tides, residual currents, and winds. The recommended method is to continuously inject a tracer into the effluent for three tidal cycles (under spring tide conditions), and to measure the plume during the slack current after the second and third tidal cycles. Tracer concentration profiles collected in the plume in two directions set at right angles are recommended. The depth of the profiles is determined by measuring the tracer concentration depth profiles in the plume, then selecting the depth of the maximum concentration as the depth for the tracer concentration profiles.

At least two recording current meters at the maximum concentration depth and a wind anemometer are recommended during the tracer studies and for at least a 60-day period after the studies in the vicinity of the discharge.

From the tracer profiles, the effluent plume envelopes are developed and are related to the measured currents. The probability of the occurrence of these envelopes is determined from statistical analysis of the current meter records.

Combination Exposure Areas

Exposure areas can sometimes be comprised of several different types of water bodies, for example, a series of lakes and streams. In these situations, plume dispersion models may not be able to accommodate great variability in downstream conditions. Such scenarios may call for the formulation of site-specific models. Downstream water quality relationships to be considered in the development of simple, site-specific dilution models can be found in relevant reference material (e.g., Thomann and Mueller 1987; McCutcheon and French 1989; Henderson-Sellers and French 1991; Lung 1993; and Tchobanoglous and Schroeder 1987). Should analysis require the application of more complex models, a wide variety of sophisticated modeling packages are readily available as shareware tools or at a minimal cost to the user.

3.3.3 Using Predictive Models

The use of predictive effluent plume models allows for a more complete characterization of the effluent plume based on samples collected during plume delineation studies. Modelling provides a means of extrapolating from the plume measurements to define the extent of far-field conditions and to simulate the continuous injection of effluent typical of site operating conditions. Plume models also assist in the interpretation of collected data when the plume configuration is strongly influenced by the direction of ambient currents (e.g. lakes and oceans) and when the exposure area is characterized by a long flushing time (e.g. stagnant waters). It is recommended that, where possible, predictive models be verified with some field measurements.

When ambient currents modify plume configuration, the plume's characteristics can be established through statistical analysis of the data collected by the current meters. For conditions of a long flushing period, the measured dye concentrations are used to calibrate a numerical transport-diffusion model. The model can then be used to simulate the effluent delineation and characteristics arising from a continuous discharge. The model can be run for a variety of conditions (e.g. seasonal variations of water movements and wind patterns), thereby overcoming the limitations of the particular conditions recorded in a single field study. The output from the model can be statistically analyzed to estimate the desired regions of maximum extent, average conditions, and minimum dilution. Discussion of appropriate plume models capable of such calculations can be found in U.S. Army Corps of Engineers (1986), and in Liu and Leendertse (1987).

3.4 Use of Sublethal Toxicity for Site Characterization

Since historical sublethal toxicity data for some or all of the required tests (see Section 7.1) has been generated for effluents discharged from a number of mine sites across Canada, the operator may want to use this information during site characterization in the following ways:

1) To aid in determining sampling areas for fish or the benthic invertebrate community. If the operator has no historical field data on fish, fish food sources or fish habitat from their effluent exposure zone, then historical effluent sublethal toxicity data (if available in sufficient quality) could be used to estimate the potential zone of influence to help in establishing sample collection locations for the fish and benthic invertebrate community surveys in Initial Monitoring (ie: estimation of extent of response in the near field zone). Details on how to estimate the geographic extent of a sublethal toxicity response are provided in Chapter 7.

2) As a tool to estimate effluent concentrations in the case of confounding factors. If there are numerous effluent discharge locations at a mine site, the Microtox Chronic test (see Section 7.5.1) or one of the other recommended sublethal toxicity tests could be used to quantify the degree of sublethal toxicity contributed from the different effluent discharge sources as an aid to determining whether or not an effluent is considered a significant discharge.

3.5 Characteristics of Mining Environments

Many mines and mining activities share some similarity in environmental features. They are briefly outlined below:

3.5.1 Headwater Location of Mine Sites

Many mine sites are located near the headwaters of rivers or streams. In some instances the effluent can constitute a significant portion of the flow downstream of the discharge point. This will have an impact on how to characterize the mine's exposure area.

Headwater streams, because of their size, gradient, or intermittent flow, are often not suitable fish habitat. Therefore, the mine effluent is frequently discharged to receiving waters with small or no fish populations, although the diluted effluent will, in most cases, eventually reach fish habitat. Nonetheless, some fish species often utilize accessible headwater areas at some stage of their lifecycle, such as for spawning. The location of sampling sites, especially for fish, should take this aspect into consideration when designing the EEM study.

3.5.2 Size and Life-Spans of Mines

Mines are often limited, both in terms of affected area and effluent volumes released. They also have variable life spans, from very short (2-4 years) to very long. Mines may obtain ore from underground or open pits. Both methods involve water table manipulation, waste rock disposal, and tailings disposal. Each of these has the potential to disrupt fish, fish habitat and the use of fisheries resources.

3.5.3 Effluent Quality and Volume

Mine effluent is mainly composed of water, process chemicals and trace metals. Metals are natural substances that cannot be broken down; they are the main constituents of the periodic table and many are required in small to trace amounts to sustain life. However, many metals can be toxic to aquatic organisms at low concentrations. Process chemicals, such as ammonia and cyanide are also known to cause aquatic effects at certain concentrations.

Effluent quality and quantity is influenced by several factors including the nature of the ore and host rock, processing methods, effluent treatment methods, climate and site hydrology. In Canada, many mining operations discharge seasonally. Reduced natural degradation of substances such as cyanide and ammonia under cold conditions sometimes makes it difficult to achieve discharge limits. Therefore, wastewater is discharged during spring and summer. Wastewater may also be discharged during early spring to release large quantities of snow meltwater that has accumulated during the winter months. Other operations have relatively constant effluent discharge rates.

The diversity of metal mines and their receiving environments results in a wide variety of discharges. Discharge rates will vary both in volume and duration based on site specific factors.

In addition, the need to protect over-wintering fish in pools when natural stream flow is minimal may cause a mine to reduce its effluent discharge to the exposure area. Other mines minimize discharges in late summer when stream flow is low and fish may be spawning. Conversely, spring discharge rates may be high, allowing water inventories to be relieved while stream flows are high. However, not all mines have sufficient reservoir capacity to optimize effluent releases. These factors will therefore be important to consider when designing the EEM study.

The retention time in tailings ponds can also affect the composition of the effluent. For example, cyanide breakdown and settlement of particulate matter are time-dependent. A sufficiently long residence time can potentially modify the concentrations in the effluent.

Mine ore bodies are variable, resulting in variations in effluent discharges. Each ore is different, not only between mines but also varies within each mine itself. One ore body may contain more or less of certain minerals. In addition, while mines may have a short life span, the mill at the mine site may process ore from several mines. This would result in different quality effluents discharged over many years. Also, the acid generating potential in ores can vary between mines and within the same mine. Virtually every mine has its own distinct suite of parameters of concern in effluents that influence site specific factors related to bioavailability and hardness. However, similarities in operations among mines in a region may allow for efficiencies to be realized by a regional EEM study design.

3.5.4 Local Geology

Mine location is determined by the regional geology, and the exact locations of ore deposits. Local mineralization around ore deposits influences the natural background levels of metals in streams. The net result is often a naturally higher background concentration of metals in streams located near mine sites. Selection of appropriate reference areas during EEM has to take this site-specificity into account to increase its comparability to the exposure area.

3.5.5 Bioavailability of Metals

The effects of soluble metals on the biota can be mitigated during effluent treatment by the adding lime to precipitate metals. However, the bioavailability of metals and the modifying factors that can change this availability are still at the research stage and are the focus of many studies. The bioavailability of metals is an important aspect to consider when assessing the effects of mine effluent in the EEM program. In addition, this characteristic implies that the effect mine effluent and associated sediments can have on the aquatic environment can vary during the mine's life cycle.

3.6 References

Busch WDN and Sly PG (eds.) 1992. The development of an aquatic habitat classification system for lakes. CRC Press. Boca Raton. 225 pp

COSEWIC (Committee on the Status of Endangered Wildlife in Canada). 1998. Species presently registered as rare, threatened or endangered in Canada. http://www.cosewic.gc.ca/, COSEWIC, Canadian Wildlife Services.

Cowardin LM, Carter V, Golet FC and LaRoe ET. 1979. Classification of Wetlands and Deepwater Habitats of the United States. U.S. Fish and Wildlife Service. FWS/OBS-79/31.
Department of Fisheries and Oceans. 1990. Coastal/Estuarine Habitat Description and Assessment Manual. Part II. Habitat description procedures. Prepared by G.L. Williams and Associates Ltd., Coquitlam, B.C.

Department of Fisheries and Oceans and British Columbia Ministry of the Environment and Parks. 1987. Fish Habitat Inventory and Information Program: Stream Survey Field Guide.

Douglas GR, Grant CE, Grant RDL, Salamone MF and Nessman ER. 1983. Comparative mammalian in vitro and in vivo studies on mutagenic activity of Rhodamine WT. Mutat. Res. 118: 117-125.

Fischer HB, List EJ, Koh RCY, Imberger J and Brooks NH. Mixing in Inland and Coastal Waters, Academic Press Inc., San Diego, California, 1979.

Freeze RA and Cherry JA. 1979. Groundwater. Prentice-Hall. Englewood Cliffs, New Jersey.

Haan CT, Johnson HP, Brakenseik DL. 1982. Hydrologic Modeling of Small Watersheds. American Society of Agricultural Engineers, Michigan. pp. 135-174.

Haan CT, Barfield BJ and Hayes JC. 1994. Design hydrology and sedimentology for small catchments. Academic Press. New York.

Liu SK and Leendertse JJ. 1987. Modelling the Alaskan Continental Shelf Waters. The Rand Corporation, 1700 Main Street, P.O.Box 2138, Santa Monica, California. R-3567-NOAA/RC.

Ontario Ministry of Natural Resources. 1989. Manual of Instructions. Aquatic Habitat Inventory Surveys. MNR, Toronto, Ontario.

Orth DJ. 1989. Aquatic habitat measurements. In: Fisheries Techniques. Neilson LA and Johnson DL (eds.), pp. 61-84. Am. Fish. Soc., Bethesda, Maryland.

Parker GG. 1973. Tests of Rhodomine WT dye for toxicity to oysters and fish. J. Res. U.S. Geological Survey 1(4): 499.

Plafkin JL, Barbour MT, Porter KD, Gross SK and Hughs RM. 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and fish. EPA/444/4-89-001.

Thomann RV and Mueller JA. 1987. Principles of Surface Water Quality Modeling and Control, Harper Collins Publishers, Inc. New York, New York.

U.S. Army Corps of Engineers. 1986. CE-QUAL-W2: a Numerical Two-Dimensional, Laterally Averaged Model of Hydrodynamics and Water Quality; User's Manual. Instruction Report E-86-5, U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi.
U.S. EPA. 1987. The Enhanced Stream Water Quality Models QUAL2E and QUAL2E - UNCAS: Documentation and User Manual. Environmental Research Laboratory, United States, Athens, Georgia. EPA/600/3-87/007.

U.S. EPA. 1987. The Enhanced Stream Water Quality Models QUAL2E and QUAL2E - UNCAS: Documentation and User Manual. Environmental Research Laboratory, United States, Athens, Georgia. EPA/600/3-87/007.

U.S. EPA. 1988. WASP-4 a Hydrodynamic and Water Quality Model-theory, User's Manual and Programmer's Guide. Environmental Protection Agency, Athens, Georgia. EPA/600/3-87/039.

U.S. Geological Survey. 1986. Fluorometric Procedures for Dye Tracing. U.S. Geological Survey, Federal Center, Box 25425, Denver. Book 3. Applications of Hydraulics. Chapter A12.

U.S. Geological Survey. 1989. Determination of Stream Reaeration Coefficients by Use of Tracers. U.S. Geological Survey, Federal Center, Box 25425, Denver. Book 3. Applications of Hydraulics. Chapter A18.

Viessman W, Lewis GL and Knapp JW. 1989. Introduction to Hydrology. Third Edition. Harper and Row, New York. 1989.

Table of Contents

4. Fish Monitoring *4-1*

4.1 *Overview* *4-1*

4.2 Study Design *4-3*
4.2.1 Design Considerations 4-3
4.2.2 Selection of Exposure and Reference Areas 4-3
4.2.2.1 Exposure Area 4-5
4.2.2.2 Reference Area 4-6
4.2.3 Aspects to Consider When Selecting the Exposure and Reference Area 4-6
4.2.3.1 Tributaries and Other Point and Non-Point Source Discharge 4-7
4.2.3.2 Natural Environmental/Habitat Variation 4-8
4.2.3.3 Historical Damage 4-8

4.3 First EEM Study *4-9*
4.3.1 Review of Historical Data 4-9
4.3.2 Initial Monitoring 4-13
4.3.2.1 Introduction 4-13
4.3.2.2 Fish Population Survey 4-13
4.3.2.2.1 Selection of Fish Species For Monitoring 4-14
4.3.2.2.2 Fish Exposure 4-19
4.3.2.2.3 Sample Sizes and Role of Effect Size 4-21
4.3.2.2.4 Timing of Sampling 4-24
4.3.2.3 Fish Usability 4-25
4.3.2.4 Alternative Methods 4-26
4.3.2.4.1 Caged Bivalves 4-27
4.3.2.4.2 Mesocosms or On Site Bioassays 4-27
4.3.2.4.3 Expanded Sublethal Toxicity Testing 4-28
4.3.3 Marine Discharges 4-28

4.4 Data Assessment and Interpretation *4-29*
4.4.1 Statistical Guidance for the Fish Survey 4-31
4.4.2 Size-at-Age 4-33
4.4.3 Fecundity, Gonad Weight, Liver Weight and Condition 4-34
4.4.4 Mean Age 4-34
4.4.5 Physical Abnormalities 5.35
4.4.6 Order of Data Analysis 4-35
4.4.7 Data QA/QC and Analysis 4-36
4.4.7.1 Data Errors and Outliers 4-36
4.4.7.2 Validity of Statistical Assumptions 4-37
4.4.7.3 Independence (pseudoreplication) 4-37
4.4.7.4 Normality and equality of variances 4-38
4.5 Subsequent EEM Studies 4-38
4.5.1 Periodic Monitoring 4-38
4.5.1.1 Confirmation Monitoring 4-39

4.5.1.2 Surveillance Monitoring 4-39
4.5.1.3 Minimal Monitoring 4-40
4.5.2 Focussed Monitoring 4-40
4.5.3 Investigation of Cause 4-41
4.5.3.1 Ecoepidemiological Criteria for Attributing Cause 4-41
4.5.3.2 Identification of Specific Stressors Primarily Responsible for Changes4-42
4.5.3.3 Identification of Specific Chemicals Associated With Changes 4-42

4.6 Methods of Sampling 4-43
4.6.1 Gillnets 4-43
4.6.2 By-Catch 4-43
4.6.3 Electrofishing 4-44
4.6.4 Trawling 4-44
4.6.5 Beach Seines, Fyke Nets, Trap Nets, Hoop Nets and Fish Traps 4-44
4.6.6 Hook and Line and Hand Collections 4-45
4.6.7 Purse Seines and Midwater Trawls 4-45
4.6.8 Remote Sensing 4-45
4.6.9 Determination of Sampling Effort 4-46

4.7 ***Sampling and Measurement Precision 4-47***

4.8 ***External Appearance 4-49***

4.9 ***Special Considerations for Small Bodied Fish Species 4-50***

4.10 Fish Usability / Exposure Method 4-51
4.10.1 Fish Tissue Analysis 4.51
4.10.1.1 Tissue Sample Collection and Preparation 4-52
4.10.1.2 Supporting analyses : Lipid and Moisture % 4-52
4.10.2 Tainting Evaluation 4-53

4.11 Quality Assurance and Quality Control 4-54
4.11.1 Standard Operation Protocols (SOPs) 4-55
4.11.2 Quality Control in the Field 4-56
4.11.3 Quality Control in the Lab 4-57
4.11.4 Data Analysis 4-58
4.11.5 Data treatment 4-59

4.12 ***References 4-60***

List of Tables

Table 4-1: Guidance for the evaluation of historical studies 4-12
Table 4-2: Monitoring species characteristics for optimising the assessment of aquatic environmental health versus human health issues using fish populations 4-17
Table 4-3: Recommended fish survey measurements to determine effects in fish growth, reproduction, condition and survival 4-24
Table 4-4a: Effects on fish 4-30
Table 4-4b: Endpoints to be used for supporting analyses 4-30
Table 4-5: Sample Data Checklist for Fish Collections including Physical Abnormalities 4-49
Table 4-6: Anatomical structures commonly used for age determination (from Environment Canada, 1998) 4-50
Table 4-7: Example of quality assurance checklist for aquatic biology studies (check or enter footnote number to provide details or to describe deficiencies.) 4-57

List of Figures

Figure 4-1: Use of historical information and sequence of events leading to the completion of Initial Monitoring** 4-11
Figure 4-2: Features of composite fish 4-48

4. Fish Monitoring

4.1 Overview

Fish monitoring for the EEM program consists of monitoring the fish to determine if there are differences in the growth, reproduction, survival or condition of fish population and monitoring levels of mercury in fish tissue (as an indicator of fish usability), in order to determine whether or not the mine effluent is having an effect on fish.

Mines are not required to conduct a fish survey in situations where the effluent concentration within 250 meters from each final point of discharge is less than 1%. Methods to determine the concentration of effluent, including conducting a plume delineation study, can be found in Chapter 3, Section 3.3. The mine will report in the study design how the effluent concentration was determined.

Mines do not have to conduct a fish tissue analysis if levels of mercury in the effluent are less than 0.05μg/L, as determined during Effluent Characterization (Chapter 6).

The mine site can be characterized using historical information (Site Characterization - Chapter 3). Any missing site characterization information will be obtained by the mine, and there is the possibility that some field work may be necessary at this phase. Site Characterization information is used to design an appropriate EEM study.

As the EEM program is under the authority of the *Fisheries Act,* the definition of fish means fish as defined in Section 2 of the *Fisheries Act,* but does not include parts of fish, parts of shellfish, parts of crustaceans, or parts of marine animals. The fish survey was originally designed for finfish, and these remain the recommended fish species for monitoring. But, in some situations where monitoring finfish may not be appropriate, shellfish and crustaceans may be considered.

The first biological monitoring phase is called *Initial Monitoring*. The frequency of subsequent biological monitoring phases is dependent up the results of the previous monitoring phase. In the case where there are no effects on fish, the effects on fish have not been confirmed or there are effects on fish for which the magnitude, extent and cause are known, *Periodic Monitoring* is done to confirm previous results and to ensure that there are no changes .

If an effect on fish is found, its magnitude and geographical extent will be evaluated through *Focussed Monitoring*, the scope of which will depend on the results of previous studies.

Once the magnitude and geographical extent of the effect on fish is known, and previous monitoring has not determined the cause of the effect, identification of the cause of the effect is determined through *Investigation of Cause*.

Once the cause of the effect is known, the mine may proceed to study possible corrective actions or potential research work, which is outside of the EEM program. Environment Canada, in consultation with Fisheries and Oceans Canada, has developed a "Guidance for Determining Follow-up Actions when Effects Have Been Identified in Environmental Effects Monitoring (EEM)". This document provides guidance to regulatory agencies, regulated facilities, and EEM practitioners for determining follow-up actions when effects have been identified in EEM. It discusses the factors that need to be considered for determining follow-up actions, and explains the roles and responsibilities of government, industry, and stakeholders in this process.

At any time during the program, if an effect is seen, the mine can proceed directly to Focussed Monitoring or Investigation of Cause.

Data Assessment and Interpretation follows each monitoring phase. In Data Assessment and Interpretation, the following questions are answered:

- is there an effect?
- is the effect mine related?
- is the magnitude and extent of the effect known? and,
- is the mine-related cause of the effect known?

Following Data Assessment and Interpretation, the mines will submit an interpretative report to the Authorization Officer (AO). This report will include information such as a description of any changes from the submitted study design, the locations of the sampling areas, the identification of any effects, the conclusions of the study, and a timeline outlining when the next monitoring phase will be conducted.

If more than one mine discharges in the same area, joint EEM studies are encouraged. It is important to note that collaborative studies between mines need not mean that there is reduced sampling and a single reference area. The strength of interpretation will increase if the pooled resources are used to sample additional reference areas.

4.2 Study Design

4.2.1 Design Considerations

Study designs describe the following:

- information related to Site Characterization;
- the objectives of the field monitoring program, including overall approach and rationale for fish monitoring, which may be based on previous monitoring results;
- statistical design criteria, hypotheses, statistical methods and data needs;
- a description of how the fish survey and fish tissue analysis will be conducted in order to determine if there are effects, taking confounding influences into consideration;
- field sampling plans, including what will be measured where and when;
- QA/QC provisions in study designs;
- schedules for field monitoring and submission of the interpretative report; and
- a summary of previous biological monitoring and effluent and water quality monitoring.

Mines are known to discharge effluent at irregular rates during the year. Sampling when effluent is not being discharged for long periods should be avoided, as the timing of sampling and the choice of the fish species should be made according to normal operation of the mining facility to ensure that the effluent is present and its potential impact on the fish can be evaluated.

Environment Canada (1998) provides a description of study design and methods for conducting fish surveys. Shuter (1990) and Sandström (1996) provide general descriptions of study designs and methods. Munkittrick (1992) and Munkittrick and McMaster (1999) also review study design considerations.

4.2.2 Selection of Exposure and Reference Areas

The selection of the sampling areas is one of the most critical components of the study design and should be considered carefully to maximize the quality of the information gained from the study. At a minimum, sampling should be conducted at no less than one reference area and one exposure area during the first EEM study (Initial Monitoring), and subsequent EEM studies (Periodic Monitoring and Focussed Monitoring). It is recognized that sampling one reference and one exposure area at some mines could be difficult due to problems related to dissimilar habitats, confounding influences or site access.

Choosing a reference area as similar as possible to the exposure area is essential in order to interpret differences between areas. Although the use of a single reference area is the current recommendation for EEM, the use of multiple reference areas can strengthen the interpretation of differences (Munkittrick et al. 2000). The multiple reference area strategy should be considered because it offers the greatest statistical power to detect a meaningful difference between a reference area and an exposure area (Foran and Ferenc, 1999). The use of multiple reference areas can also give an indication of variability among reference areas (Munkittrick et al. 2000). Differences that are found in the exposure area that are outside the range of values seen at a number of reference areas is more likely to be ecologically relevant (Munkittrick et al. 2000). Sampling multiple reference areas is be preferred over increasing sample size (i.e. number of fish) at a single area (Environment Canada 1997).

The exposure area should be selected to ensure that the fish collected have been exposed to the effluent. The exposure area should be sampled first to determine which fish species are present, and their relative abundance within the area. The reference area can then be selected ensuring fish of the same species and similar relative abundance are present in order to compare fish among areas.

If there are no fish found in the reference area, it cannot reasonably be expected that fish would be found in the exposure area. If fish are found in the reference area, but none are found in the exposure area where you could reasonably expect to find fish (e.g., fish were historically found in the sampling area), the absence of fish in the exposure area should be reported as an effect. To determine if the current effluent discharge is causing effects, a mesocosm study could be conducted. For example, if there are physical barriers to fish movement in the exposure area, fish may not be able to move back into an area where there were historically fish. Therefore, the current discharge may not have the potential to cause effects, but historical practices harmed the fish populations.

Both sampling areas should be:

- as similar as possible except for exposure to effluent. It is recognized that the two areas are unlikely to be identical, however the assumption is that the differences in natural characteristics will be small relative to the potential effect associated with the presence of effluent;
- situated as closely as possible to each other (but sufficiently distant to be confident fish from the reference area are not exposed to effluent);
- sampled at the same time period, appropriate to record fish measurements;
- described in as much detail as possible, including the latitude and longitude and a written description of the area (physical, chemical and biological habitat, including measurements of temperature, depth and flow are recommended).

In selecting the sampling areas, the mine should take into consideration:

- the location of sampling areas in previous fish monitoring;
- the location of confounding influences;

- the number of fish needed to be collected, in the exposure and reference area, to meet fish sample size recommendations;
- habitat type;
- site access; and,
- other issues that could affect the mobility of fish.

The habitat for exposure and reference areas should be similar in terms of depth, substrate, flow, cover and water quality (other than the mine-related factors). It is also essential that the site be accessible and it is possible to safely sample fish during the most appropriate season (i.e., when measurements on fish growth, reproduction, condition and survival can be taken).

In cases where the mine has more than one final discharge point, it is recommended that sampling for the fish survey be done in an exposure area where the effluent has the greatest potential to have an adverse effect on the receiving environment. To determine this, the mass loadings of the deleterious substances in the effluent should be taken into consideration, the manner in which the effluent mixes in the exposure area, and the sensitivity of the receiving environment.

One factor in data interpretation will be related to the certainty that exposure area fish and reference area fish populations are not mixing. In cases where there are obvious physical barriers (e.g., dams, waterfalls, beaver dams) separating areas, there will be a high level of confidence that fish are not moving between areas. One of the most difficult situations to interpret arises when there are no statistical differences between areas, and there are no barriers restricting movement. In these cases, an indicator of exposure to the effluent is recommended (Section 4.3.2.2.2).

Where possible, sampling areas for other components (benthic invertebrates, water and sediment) should be co-located. The characteristics of the selected fish species, their mobility, habitat usage and the different sampling gear may not always make this practical. The reference areas for benthic invertebrate sampling in some cases can be directly upstream of the exposure area, which may not be the case with fish (due to mobility), so this may not always be practical.

4.2.2.1 Exposure Area

The exposure area, as defined in Schedule 5 of the MMER, is "all fish habitat and waters frequented by fish that are exposed to effluent". During Initial and Periodic Monitoring, sampling in the exposure area should be done in an area closest to the effluent discharge, with the appropriate habitat for fish, where fish are present and where one could reasonably expect to find effects. The exposure area may extend through a number of receiving environments and contain a variety of habitat types. In Focussed Monitoring, additional sampling further away from the effluent discharge will be done to determine the magnitude and geographical extent of the mine-related effect.

4.2.2.2 Reference Area

The reference area, as defined in Schedule 5 of the MMER, is "water frequented by fish that is not exposed to effluent that has fish habitat that, as far as practicable, is most similar to that of the exposure area". The reference area is ideally located immediately adjacent to or upstream of the exposure area and should be sampled as soon as possible after the exposure area. The reference area should not be exposed to effluent from the mine under study and should have similar hydrology and habitat as the exposure area.

Baseline data (pre-effluent discharge), and multiple reference areas may assist in data interpretation. When a reference area is not used and a mine only uses pre-discharge data to determine effects, any changes from this data in comparison with the exposure area will be attributed to the mine.

It is possible to use historical data as a baseline comparison to determine effects. In situations where there are historical data available for comparison, it is recommended a reference area still be sampled, since there can be changes in fish parameters related to changes in environmental conditions (e.g. due to flooding or variability in annual temperatures). A reference area should be used to allow characterization of those changes that are mine-related from those which are not.

It should be noted that suitable reference areas may not be available adjacent to or immediately upstream of some mines. For example, there may be dams and reservoirs immediately upstream of mines on rivers. In those cases, other possibilities can be explored, such as using a reference area located further from the discharge point or sampling a number of exposure areas at increasing distances from the point of discharge, representing an exposure gradient (gradient design). These approaches would require either well separated areas for larger species, or the use of small fish species. Barriers to fish movement which have not resulted in disruption of habitat (e.g., waterfalls) and for which suitable reference areas are available may aid in data interpretation.

4.2.3 Aspects to Consider When Selecting the Exposure and Reference Area

Potential problems when selecting reference and exposure areas are:

1. mobility of species (fish in an upstream reference area could have been exposed previously; fish in the exposure area could be transient, reducing exposure to potential effects);
2. natural variability;
3. differences in habitat; and
4. confounding influences.

Variability in fish parameters (e.g. growth, reproduction, condition, survival) are a function of numerous factors, not all of which will be related to the effluent. The fish survey attempts to reduce variability through stratification of sample timing and habitat type. The selection of appropriate fish species for monitoring, survey timing and sampling gear will also facilitate the interpretation of any differences detected. Nevertheless, natural and other anthropogenic factors may influence effects on the fish and fish tissue and confound interpretation of data.

The fish survey attempts to guard against temporal differences through the use of confirmation monitoring (i.e., the same effect found in two consecutive phases of monitoring) when effects are detected, prior to determining the magnitude and geographic extent of the effect (Focussed Monitoring). Incorporating multiple reference locations into the study design can aid in designing against spatial confounding factors, and practitioners are encouraged to do so. Design considerations for the detection of anthropogenic disturbances have been presented in the literature (see Green 1993 and references therein; Underwood 1994; Underwood 1997), and practitioners are encouraged to incorporate these considerations into their study designs.

Some examples of potential confounding factors include:

- tributaries and other point and non-point source discharges (e.g. other industrial discharges, agricultural run-off, aquaculture facilities, sewage treatment plants);
- natural environmental/habitat variables;
- historical damage.

Attention to potential confounding factors identified during Site Characterization (Chapter 3) should be considered during study design. In this way confounding factors can be minimized, or accounted for in the study design so that their influence can be assessed during data interpretation.

4.2.3.1 Tributaries and Other Point and Non-Point Source Discharge

Tributaries provide dilution water to a main river channel, lake, estuary or ocean. The tributary water may or may not have similar chemical properties to the water body under study. It may also take some distance for a tributary to mix with the water body. If there is a tributary between the reference and exposure area, the additional flow from the tributary can potentially confound the interpretation of data.

Other point and non-point source discharges may make it difficult to distinguish between effects due to the mine effluent and other discharges, particularly if they are found within close proximity to the mine discharge. When other discharges are present immediately above the effluent, multiple reference areas should be used. One reference area should be set between the other discharge and the effluent discharge. In this way it may be possible to account for the influence of the other discharge. As well, the reference area should have similar background levels of metals.

If there are other point source effluent discharges not related to mines in the study area, the study design should attempt to minimize the potential effects of the confounding factors and potential effects from the mine. When it is not possible to resolve the confounding factors by modifying the study design, alternative sampling designs and methods should be evaluated.

4.2.3.2 Natural Environmental/Habitat Variation

Natural biological communities can differ both temporally and spatially. The influences of effluent, if any, relative to natural variation, may be difficult to determine.

Examples of common, naturally occurring confounding factors to consider include:

- habitat type (riffle, run, pool);
- substrate type (organic content, particle size);
- water depth;
- water flow rate/discharge;
- tidal action/currents/wave exposure;
- salinity;
- dissolved oxygen/temperature;
- emergent/submergent vegetation cover;
- water chemistry (conductivity, hardness, pH, etc.,); and,
- biological properties.

4.2.3.3 Historical Damage

When the area in which a mine releases its effluent has been subject to previous damage from other activities, it may be difficult to determine differences due to current effluent release practices.

4.3 First EEM Study

4.3.1 Review of Historical Data

Prior to submitting the first EEM study design, a review of historical information or data collected by another regulatory agency should be undertaken. Historical data can be submitted as part of the study design, as it may provide useful information for site characterization. This data may also be used to determine if the effluent is causing an effect on the fish population, fish tissue **or** the benthic invertebrate community. Figure 4.1 provides guidance on how historical data may be used to proceed to Data Assessment and Interpretation.

If the mine can determine whether or not there are effects on the fish population, fish tissue or benthic invertebrate community, the mine may submit a report to the AO within twelve months of the mine being subject to Section 7 of the MMER, that outlines the scientific data and results and interpretation of the historical study. The mine would then submit the first EEM study design 24 months after the mine becomes subject to Section 7 of the MMER.

At a minimum, historical fish data should be able to document effects in growth, reproduction (or larval abundance, gonad development), age distribution of one or two fish species. A comparison of condition (condition factor or liver size) and physical abnormalities of reference and exposed populations should also be available.

The report outlining the results of the historical data that is submitted to the AO should include:

- confirmation that the samples collected were representative of the sampling areas at the time of field sampling;
- confirmation that the conditions during previous sampling were representative of current conditions (e.g. effluent quality) ;
- sample size for the fish survey;
- a description of the habitat type at a level of resolution sufficient to choose fish sampling areas;
- documentation of the existence of any potentially confounding or influencing factors to be considered in the study design and interpretation of results;
- a description of the fish (and shellfish) community present in the study areas;
- a description of the quality and use of fisheries resources in the exposure and reference area; and,
- a description of the sampling areas (exposure and reference areas).

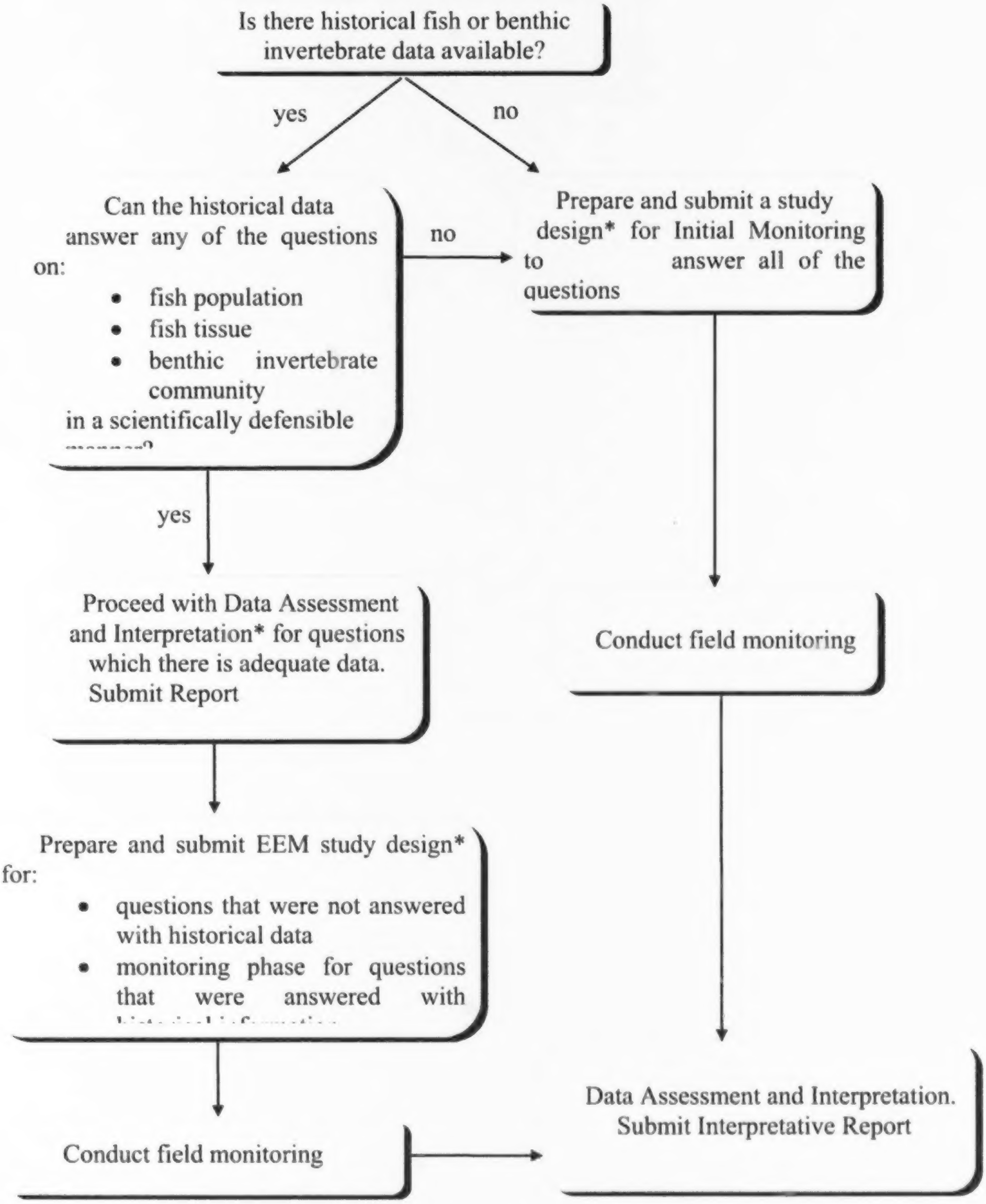
In addition to being scientifically defensible, historical data should be less than six (6) years old.

If the mine operations or the environment in the study area have changed since the last survey, the use of historical data for the first EEM study is not recommended. Any event that has had the potential to modify the effects of the effluent on fish should be taken into consideration.

Examples of such events are:
- processing of ore of different chemistry;
- mining of new ore body of different chemistry;
- change in process or treatment of effluent;
- flooding of the receiving area, affecting either or both reference and exposure areas;
- change in the location of planned final points of discharge; or
- presence of new confounding factor.

The AQUAMIN process (AQUAMIN Working Group 1, 1994) used a detailed set of criteria to evaluate historical information (pertinent items are summarized in Table 4.1). The following criteria should be taken into consideration to assess the adequacy of historical studies to be used within the fish survey component of the EEM program:
- did the study design meet its stated objectives?;
- did the study include a comparison of an exposure area to a reference area, temporal comparisons are also valid (i.e. before versus after discharge)?;
- did the study provide information regarding the methods used?;
- is information available on the quality control of the analytical and field data?;
- is information available on the statistics used in the analysis of data?

Figure 4-1: Use of historical information and sequence of events leading to the completion of Initial Monitoring**



Is there historical fish or benthic invertebrate data available?

Can the historical data answer any of the questions on:
- fish population
- fish tissue
- benthic invertebrate community

in a scientifically defensible

Prepare and submit a study design* for Initial Monitoring to answer all of the questions

Proceed with Data Assessment and Interpretation* for questions which there is adequate data. Submit Report

Conduct field monitoring

Prepare and submit EEM study design* for:
- questions that were not answered with historical data
- monitoring phase for questions that were answered with

Conduct field monitoring

Data Assessment and Interpretation. Submit Interpretative Report

**** when preparing the study design and interpreting data, the mine should take into consideration any confounding factors in the sampling area which may preclude data interpretation***

Table 4-1: Guidance for the evaluation of historical studies.

Sub-criteria	Points to consider
	Study Design
Within-station variation AND	Site-specific variability considered by replication of fishing stations.
Number and location of sampling stations AND	of location and number of fishing stations is based on: • exposure to effluent or • expected fate of contaminants or • • pre-sampling or • • another approach to optimise the design (sampling strategy: stratified random, systematic; statistical power, environmental variability and any factors affecting sensitivity)
Related factors or conditions	Relevant environmental factors are reported. Supporting information should include (when relevant) and is not exclusive: → water: pH, hardness, conductivity → biota: habitat description Confounding factors such as other sources of pollution are identified
	Reference Areas
Spatial reference station	Reference area(s) located where it is not exposed to effluent (e.g. upstream or beyond any influence of the discharge), in the same water-body if possible or in another water-body. environmental conditions in reference and exposure areas are comparable (similar natural habitat features).
	Methods
Standardized methods OR	Methodologies used are based on validated procedures prepared by recognised government/agency, national or international standard-setting organisations such as ASTM (American Society for Testing and Materials), USEPA (United States Environmental Protection Agency), Environment Canada, OECD (Organisation for Economic Co-operation and Development), WHO (World Health Organisation).
Referenced methods OR	Methods used are in the published literature and accepted as adequate protocols.
Described methods	Detailed written methods are provided. Method description should be very detailed (sampling equipment, sample collection, holding, storage, sample preparation, etc.). Field collection methods are documented for achieving and maintaining sample integrity. Novel approaches are well documented.
	Quality Control
Detection limits	Detection limits are provided or can be easily calculated/determined from data. Detection limits for the different types of studies include; for biological studies - level of identification (taxonomic levels defined) or identification keys are referenced; for chemical analysis - method detection limit.
Precision	Precision is expressed as standard deviation or coefficient of variation. For chemical data, the analytical variation among laboratory replicates (split samples) is reported. For tissue metal levels data, duplicate processing is used.
Accuracy	Accuracy is expressed as a percentage of the deviation from the true value and should be evaluated using recognised relevant best available reference material. For biological data, reference collections of organisms are reported. Method blanks are used to determine the level of contamination introduced into the sample during the processing and analysis.
	Data Analysis
Measures of effects	The magnitude of effects (e.g. severity or degree of the impact) is described.
Data assessment	Data analysis summarizes chemistry and biological data, or to test hypotheses of significant changes or to test the ability of the study to detect changes.

Adapted from AQUAMIN Working Group 1, 1994.

4.3.2 Initial Monitoring

If there are no historical data or the historical data are insufficient to determine if there are effects in the fish population, fish tissue **and** benthic invertebrate community, Initial Monitoring will be conducted.

4.3.2.1 Introduction

The effect of effluent on fish will be evaluated by comparing population indicators (growth, reproduction, condition and survival). Levels of mercury in fish tissue and altered flavor or odor (fish tainting) may also be evaluated to determine if there have been effects or changes in fish usability. To evaluate the effect of effluent on fish, the following questions will be answered:

- have fish been modified by the effluent?
- has there been a change in fish usability due to effluent?

Local input should be taken into consideration when designing the fish survey in order to provide information on the presence of fish species or information that may be helpful in determining sample locations.

4.3.2.2 Fish Survey

To answer the question "have the fish been modified by effluent?", effects on growth, reproduction, condition and survival of the fish population are determined. Survival (age), energy use (growth and reproduction) and energy storage (condition) are measurements that can detect an effect in the fish population related to effluent.

Growth is the change in size (weight or length) with time or age. Reproduction is expressed as reproductive effort, fecundity, egg size or gonad weight relative to body size. Reproduction may be the most sensitive measurement in resident fish. Alterations in reproductive investment can be evident within a year, since the reproductive tissue is generally turned over annually. Fecundity and gonad weight are easy to measure, if an appropriate sampling time is chosen. Monitoring should be conducted prior to the spawning season, when gonads are large and well developed. To determine appropriate timing for sampling, knowledge of the fish species life cycles are essential.

Information on relative year class strength can also be used to assess reproductive success. If non-destructive sampling is carried out, relative year class strength (young of year survival, % composition of age classes) will be the only way to determine reproductive success. In situations where early life stages/juveniles reside in the exposure area, larval density could be used to determine reproductive success.

During the fish survey, the physical state of the fish is also assessed. Condition can be used as an index of well being and should be reported. A visual estimation of physical malformations including lesions on the body surface, eroded, frayed or hemorrhagic fins, parasites or other physical deformations is recommended. If malformations are more frequent in fish from the exposure area, a more detailed investigation can be carried out during a subsequent phase of the program (Focussed Monitoring or Investigation of Cause).

4.3.2.2.1 Selection of Fish Species For Monitoring

Community Survey

Traditional fish community surveys evaluate whether there are differences among areas in the variety and abundance of fish species present. Information on the relative abundance and presence or absence of species will also aid in the selection of appropriate fish species (Section 4.3.2.2.1) for the fish and usability surveys and for the design of subsequent phases of the program. Fish communities often include a number of species that are not abundant for a variety of reasons that may be unrelated to the presence of effluent. This evaluation should be based on abundant species present in the reference area and conclusions to the fish community question should not be based on species that are not abundant in the reference areaA change in fish community has occurred when species that are expected to be abundant from the collections conducted at reference areas are not present in the effluent discharge area. If the downstream area does not support one or more of the abundant species found at the reference area it will be necessary to document how far downstream this absence persists. When the fish community composition has changed because of the presence of a effluent, there will also be measurable changes in the fish populations that remain.

If a mine has sufficient historical data to determine whether the appropriate species are present and can choose acceptable fish species, and is aware of the presence or absence of fish species in the area designated as rare, endangered or threatened, the mine can then proceed to answering the fish survey objectives.

Table 3.5 lists the species presently registered as rare, threatened or endangered according to the Canadian Wildlife Service (COSEWIC 1998). This list is updated regularly and is available at the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) web site (http://www.cosewic.gc.ca/). In this list, the item "last examined" represents the date COSEWIC gave the designation to the species or updated status. Provinces and territories may have lists with designations that differ from the COSEWIC listing. For example, a species may be endangered in a province or a territory, but Canada wide, it may only be vulnerable. If this is the case, the provincial list should be used. Following the identification of the fish species for the fish survey, field sampling should be designed to limit mortality of non-target species.

In the absence of historical data, relative abundance (comparisons of catch per unit effort) or presence of fish sufficient for selecting the fish species are the minimum data to determine the appropriate fish species for the fish survey.

Non-destructive techniques are preferred for the community survey, where possible (see section 4.6 for discussion on appropriate methods).

Selection of the Fish Species for Monitoring

The most important factors when selecting fish species for the EEM program are exposure (see section 4.3.2.2.2), abundance, relevance to the study area (Munkittrick and McMaster 1999) and sensitivity to effluent. The objective of the fish survey is to evaluate the effects of effluent on the fish population and fish usability. If fish are not resident for their entire life cycle in the exposure area, then an assessment on the potential impact of the life stages of fish that are resident should be considered. If fish are only transient residents in the area, then an evaluation of whether the species that are most exposed show changes as a consequence of that exposure may be possible.

The recommended method for carrying out the fish survey is to monitor adults (sexually mature fish) of two species of relatively sedentary finfish that have been exposed to effluent over a long period. In environments that do not support adequate numbers of fish to meet the recommended sample sizes (section 4.3.2.2.3) or there are not two suitable fish species for monitoring, the following options, in orderof preference, may be considered:

- one sexually mature fish species and one sexually immature fish species;
- two sexually immature fish species;
- one sexually mature fish species; or
- on sexually immature fish species.

In situations where it has been determined that fisheries resources may be impacted by a destructive fish survey, non-destructive sampling techniques may be used.

Species that are migratory or spend only a small proportion of their life cycle in the system under investigation (e.g. anadromous salmonids, some marine fishes) are generally not suitable fish species for monitoring because exposure to effluent is minimal or transient. This is also true for species that are highly mobile and are likely to be moving in and out of the exposure area. In some cases, it may be possible to select more mobile species (e.g. mountain whitefish - Swanson 1993) for monitoring due to physical constraints which limit movement (e.g. dams, natural barriers, changes in habitat). In general, the greater the likelihood that a fish species is exposed to effluent, the greater its value in the EEM program.

Sampling abundant species in both exposure and reference areas will ensure that there are enough individuals for a successful long-term monitoring program. Sampled species should also be suitable for measuring the recommended variables. For example, fish that are difficult to age will make interpretation of results difficult (Munkittrick and

McMaster 1998). Likewise, the species role in the food web needs to be evaluated, for example some species depend mostly on terrestrial food sources for their diet. Such species would be of little relevance in assessing the aquatic impact of effluent.

Commercial, sport fish and subsistence fish populations are usually not suitable fish species for monitoring because the exploitation of their population may obscure effects related to the effluent. Large species should also be avoided, since adult abundance of small species will respond more rapidly to stressors that affect fecundity and/or larval survival (Shuter 1990). Sexually mature fin fish are preferred because measurement of reproduction parameters are easy.

Because metals can accumulate in the sediments, fish species living close to the water-sediment interface and feeding on sediment benthic invertebrates may show effects from the effluent faster. However, if biomagnification is the concern, predatory fish would be favoured. In addition demersal species (bottom dwellers) are often better suited to potential point source impact monitoring because they are lower trophic web species and will respond faster to modifications in the environment compared to higher trophic level species. Table 4.2 summarizes different characteristics a fish species should have according to the endpoint investigated. Human health and environmental health issues are presented separately because the concerns of these issues are not always similar. In most cases, effluent fall in the point source category, although in some cases, the mine property itself forms the point source.

Criteria for fish species in monitoring programs often match characteristics of intermediate-sized benthivores such as suckers. Since suckers have been successfully used in similar surveys (Munkittrick and Dixon 1989; Sandström 1996; Environment Canada 1997), intermediate-size benthivores may be suitable at many mine sites.

Smaller bodied fish species (e.g., darters, minnows, sculpins) should also be considered as potential fish species for monitoring (Gibbons et al. 1998a,b). A small bodied fish can be considered as a fish species that has a maximum size of 150 mm or less. Information on maximum growth of fish species can be found in the scientific literature, including Scott 1967; Scott and Crossman 1973; Fritz et al. 1975; Roberts 1988; Nelson and Paetz 1992; Jenkins and Burkhead 1993; Coad 1995; Leblanc and Couillard 1995; Leblanc et al. 1998. Shuter (1990) noted that small bodied fish species were more sensitive to acidification than larger species or at least responded more quickly to changes in pH.

Table 4-2: Monitoring species characteristics for optimising the assessment of aquatic environmental health *versus* human health issues using fish populations.

Characteristics	Fisheries health		Human Health	Comments
	Point Source	Non-point source		
Residency (in absence of barrier)	Local	Wide ranging	Issue-specific	For non-point source, need a species that will integrate the signals from the zone
Abundance	High	High	Issue-specific	For human health – there may be local concerns associated with food preferences and consumption which outweigh all other factors
Longevity	Short-medium	Short-medium	Long	For fisheries health issues – long-lived species will decrease the likelihood of detecting changes. For human health, long-lived increase the body burdens and consequences.
Food preference	Benthic	Issue-specific	piscivores	For human health –a species that is at the top of the food chain. For other issues, this would be associated with (usually) increased mobility
Fecundity and Growth Rate	High	High	Low	High energetic requirements are preferred, so that changes in food availability or quality will be detected quickest. For human health, retaining body burdens by slow growth and reproduction would be preferred
Age to maturation	Short	Short	Long	For impacts on fisheries health, a species which needs energy for initiation of spawning, while retaining needs for fast growing will show impacts sooner
Spawning time	Site-specific	Site-specific	Site-specific	The relationships between exposure and spawning time will vary with the issue. In prairie systems, spring spawning fish develop eggs over-winter, during maximum exposure from point-source. In other systems, maximum exposure occurs during the late summer or fall.
Food chain involvement	Yes	Yes	Yes	A species with an aquatic-based diet, and not one depending predominantly on terrestrial-based

Source: Munkittrick and McMaster, 1999.

On a practical level, small bodied fish species are usually more abundant, easier to capture, and more sedentary than larger bodied fish species. In freshwater fish, home range size has been positively correlated with body size (Minns 1995) and measurements for the fish survey can be made precisely and accurately on these small-bodied species.

There can be disadvantages to using small bodied fish in these surveys. Often, less is known about their basic biology and life history, particularly their spawning habits, making it difficult to determine the best sample areas, times and methods. Some species are multiple spawners (produce more than one clutch of mature ova every year) (e.g., Heins and Rabito 1986; Burt et al. 1988; Paine 1990). Reproductive effort in these species is difficult to estimate from a single sample, because the reproductive tissue can be turned over almost completely between clutches (i.e., most of the mass of ova in the ovary will be spawned, and then a new clutch of mature ova will be developed). The number of clutches produced during the spawning season becomes the important reproductive variable and is difficult to estimate for an individual female in the field, even with frequent sampling. It will be difficult to evaluate the significance of changes in egg production in multiple spawners if they show normal reproduction in the first clutches. Impacts will be difficult to detect and will be affected by a variety of site specific factors such as survival, immigration, emigration, etc. Until more information is known, it is assumed that multiple spawners showing normal fecundity, normal egg size, normal size and normal age at maturity prior to the first spawn of the season are not experiencing a decrease in reproduction. Alternatively, the frequency distribution of age or length classes may be used as an index of reproductive success.

Multiple or protracted spawning may also be limited in species in which the male guards the developing embryos, and in other species with large ova (Paine 1990; Gibbons et al. 1998a). As a general rule, multiple spawning will be less likely to occur in species with short spawning seasons, and those species are the preferred ones to monitor. In any case, a survey of multiple spawning species should be conducted prior to the first release of eggs.

Some areas may not be inhabited by adult finfish, but are nursery areas for their juveniles. In such cases, it is possible to replace the fish survey with a survey of juvenile fish. Relevant measures would include: abundance (young of year survival, percent composition of age classes), deformities associated with exposure to effluents such as vertebral fusions and compressions, spinal curvatures including lordosis and scoliosis, fin erosion, and growth in juveniles exposed to effluent compared to juveniles in the reference area. Methods for the collection of juvenile fish are well established and many juvenile fishes can be aged (e.g., Secor et al. 1995). The priority considerations here are residency and exposure of the juvenile fish.

Where no fin-fish are present, collection of wild molluscs, such as oysters or mussels may be considered. Molluscs are a diverse taxonomic group that include bivalves and gastropods and are found widely distributed throughout Canada. Freshwater bivalves have been successfully used for population level assessments of metal contamination (Elder and Collins 1991; MacMahon 1991).

Molluscs possess many qualities that a species for monitoring should exhibit:

- they are sedentary, although some species (i.e. unionids) may migrate short distances (meters) within their habitat;
- they are widely distributed across Canada and are identified with limited taxonomic expertise;
- limited information on the conservation status of molluscs suggests that most unionid bivalve species are currently stable in Canada;
- individual unionid bivalves are large enough to provide sufficient tissue for metal analyses but pooling of individuals of smaller sphaeriids and most gastropods would be required for tissue metal analysis;
- molluscs are relatively hardy and tolerate wide ranges of metal concentrations, but their tolerance is limited when pH levels are below 4.7-5.0 and the calcium concentration is below 2-2.5mg L^{-1}.

4.3.2.2.2 Fish Exposure

When there are no physical barriers between reference and exposure areas, and it is possible for fish to be moving between these areas, an indicator of exposure to the effluent is recommended.

It is possible to detect differences in exposure by examining metal levels in indicator tissues. The indicators and the tissues will vary with mine type and species being used. In general, gills, liver and kidney have the greatest potential for estimating exposure and bioavailability of metals. Mercury is the only element that has been found to accumulate in muscle tissue, so if mercury is a contaminant of concern, dorsal muscle tissue should be analyzed. Blood and bone tissue may reflect exposure to lead, and might be considered if lead is the primary element of concern (Hodson et al. 1984). Bone concentrations are expected to be most indicative of long term metal exposure, while blood concentrations are indicative of short term exposure (AETE 1998).

For larger species, samples of liver or kidney can be collected. The tissues should be frozen for later preparation and analysis. For small species (< 10 cm), whole body levels can be examined, or levels in the carcass after removing the digestive tract.

It is crucial that studies are designed to maximize the possibility of detecting effects if they are present. This includes sampling at the proper time of year, with appropriate gear, and that appropriate reference areas are used. It is essential that the studies are designed to maximize potential exposure. It will be impossible, in most situations, to select fish species that spend their entire life cycle within the exposure area. Many fish species undergo seasonal movements that may bring them into, or out of, contact with the effluent, including movement to and from spawning areas, feeding areas and overwintering areas. Selecting a sampling time and fish species that have life history habits that may increase their likelihood of exposure can maximize potential exposure.

The suitability of fish species should be evaluated at the end of each monitoring phase, based on the site-specific nature of the results and the site-specific concerns about residency. There are two main issues dealing with residency:
a) Whether the fish from reference and exposure areas were mixing, and
b) Whether the fish captured in the exposure area were resident.

When effects are seen in one or both fish species being monitored, it is required that confirmation of these effects be undertaken in the next monitoring phase (Periodic Monitoring - confirmation). When effects are confirmed, it will be necessary to document the magnitude and geographic extent of the effect (Focussed Monitoring). One of the objectives of the subsequent phases may be to document whether other species demonstrate a similar effect (e.g., to determine the magnitude). Study designs for Focussed Monitoring should include examining species that may have a higher probability of exposure.

If no effects are detected in both fish species in a monitoring phase, it will also be necessary to confirm the absence of effects in a subsequent phased (Periodic Monitoring - surveillance). During the study design of the subsequent phase(s), it is crucial to evaluate whether the species previously used maximized the potential for detecting effects at that site. The factors that need to be evaluated include sampling time, sample sizes, the power of the analyses, and the exposure of fish to the exposure area.

There are concerns about situations where there are no fish resident in the exposure area. This can occur for at least two reasons:
a) fish avoid the exposure area and fish would be absent or in reduced density
b) fish move freely in and out of the exposure area, and no species spend significant periods of time in the effluent.

In the second situation, fish may move in and out of the effluent either because they do not perceive the effluent to be an issue, or the conditions are sufficiently adverse that they avoid residing in the effluent for long periods of time. In either case, for the purposes of EEM, the goal of the study design is to determine whether fish collected in the exposure area demonstrate differences from the reference areas.

If fish demonstrate exposure, and they are collected in the exposure area, and they demonstrate differences from reference area fish, then there should be no controversy. Follow-up studies can examine other species to see if other species demonstrate differences.

If fish demonstrate exposure, and they are collected in the exposed area, and they do not demonstrate differences, then there seems to be some controversy. There is confusion about the goals and objectives of EEM in situations where residency is a concern.

"The objective of the EEM program is to evaluate the effects of effluent on the aquatic environment, specifically fish, fish habitat and the use of fisheries resources. This information should reveal the effectiveness of pollution prevention and control technologies, practices, programs and regulations."

There are at least three reasons that a mine in compliance with the MMER may not have receiving water effects:

a) the mine may not produce compounds capable of causing those effects,
b) the dilution at the site may be sufficient that the effects are not manifested,
c) there may be insufficient critical habitat in the exposure area, such that fish do not spend sufficient time in the effluent to demonstrate impacts (if any were possible).

It is outside the scope of EEM to determine why exposure area effects are not seen, when they are absent. If fish collected in the exposure area do not demonstrate effects, and this can be confirmed in subsequent monitoring phases, and the study design was adequate, then it would be concluded that that the conditions of the area allow for fish that are exposed to effluent not to be affected by it using the current design of the program, when the mine is in compliance.

For the fish survey, what we have effectively done is drawn a line in the sand that says:

> "If you can collect sufficient adult males and females of two species, and if you can demonstrate that those two species have an age, growth rate, condition, gonadal size and liver size, and fecundity and egg size, that is indistinguishable from comparable reference fish, we will conclude that there are currently no environmental effects associated with that mine when it is in compliance".

The government reserves the right to change the position of the "line" when sufficient research becomes available to indicate that the current design will miss situations that are a significant concern.

4.3.2.2.3 Sample Sizes and Role of Effect Size

If a mine has no background data for developing a program, the minimum sample size recommended, when conducting lethal sampling, is 20 sexually mature males and 20 sexually mature females of two fish species. If small-bodied fish species are chosen as one or both of the fish species, an additional 20 sexually immature fish should also be collected.

The rationale for using 20 fish of each sex for lethal sampling is that there is little change in the 95% confidence limits with increasing sample size beyond 20 fish. For example, Munkittrick (1992) suggested that there was little improvement in white sucker variance estimates with a sample size above 16. More information on calculating confidence limits about standard deviations is available in Environment Canada (1997), Zar (1984) and in Snedecor and Cochran (1980).

If a mine chooses non-lethal sampling of fish to evaluate the effects of effluent on the fish, a minimum of 100 individuals should be collected in each sampling area. The measurement endpoints will include length, body weight, age of fish, external condition and sex (if possible). The fish collected should represent the whole range of fish sizes and be representative of the population. The use of a minimum of 100 fish in each area will give a good idea of the population distribution when plotting endpoints such as the length-weight frequency. As well, when examining differences between the relative abundance of young versus mature fish, fairly good resolution is achieved (M. Gray, personal comm.). If a small bodied fish species is being sampled, may only have a life expectancy of 3 or 4 years, the 0+ and 1+ fish will constitute a significant portion of the population (for reference sculpin up to 50 to 70% of the population).

The strength of the study design can be described by its statistical power. The statistical power of a comparison is a function of the sample size, the variability and the difference observed between areas. It is important that sample size and variability be examined early in the design. To determine the sample size for detecting a specific difference, some knowledge is needed about the variability of the sample, and the statistical power level that is acceptable for the decision-making process. The purpose of defining an effect size and power level is for determining when the sampling program is collecting sufficient information for decisions to be made.

If there are pre-existing data, then the sample sizes can be calculated, if the power and an effect size that you wish to detect are known. There is considerable controversy surrounding the definition of an effect size. Barnthouse et al. (1989) argue that a 10% change in variables would be societally and ecologically significant, although they were concerned primarily with laboratory toxicity tests and not field surveys. Their proposed effect size was deliberately conservative (small) because of concerns about the uncertainty in extrapolating laboratory results to the field (Environment Canada, 1997). In wild fish, differences as high as 20% have been seen in parameters between reference areas. The pulp and paper EEM selected a target effect size of 25% in terms of gonad size. Since variability will also vary between sampling campaigns, the target effect size should not be a fixed number, but rather should be a range of changes that you wish to detect, such as 20 to 30% difference. Sample sizes can be calculated using methods described in Green (1989) and Environment Canada (1998). Appendix 1 of the Fish Expert Working Group report (Environment Canada 1997) provides a detailed discussion of power relationships, effect sizes and the benefits of reducing variability in terms of increasing power.

It is important to understand that the variability and power will vary with the parameter being studied. Fish are not equally variable across all of their characteristics. Reproductive variables are usually as variable, or more variable on a relative scale than other variables such as length, weight and liver weight (Environment Canada 1997). If effect sizes are also expressed on a relative scale (i.e., as % differences), then any study that can detect a ±25% difference in relative gonad size can detect similar or smaller differences in other important parameters.

It is recommended that α and β be set equal, at a minimum of 0.50. Refer to the "Data Interpretation Guidance for Environmental Effects Monitoring" (Chapter 9) for the rationale of setting α and β equal.

The most benefit in the fish survey is gained by decreasing variability. When variability is so high that sample sizes are not justifiable or cost effective, the first consideration should be to

a) reduce variability,
b) select alternative species that may be less variable, or
c) consider an alternative method.

Our recommendation is to design a program that can detect a difference of 20 to 30% in gonad size, using a power level of 0.95 (1-β). The magnitude of the difference that could be detected for other parameters would be fixed based on the sample size for determining an effect on gonad size. The power for detecting differences in other parameters should be reviewed during study design to ensure that reasonable power is achieved for as many variables as possible. Studies are designed site-specifically and the priority should be given to reducing variability rather than increasing sample sizes (see Section 4.2.2).

The same approach used to identify a target effect size for relative gonad weight should be applied to other variables. Sensitivity analyses using population models should be used to explore the consequences of the effect size chosen for any and all variables (Environment Canada 1997). When preliminary analyses show that power will be insufficient given reasonable sample sizes, the assessments should be redesigned.

Any frequency distribution approach will be more objective if sampling gear is non-selective so that as many age classes as possible are represented in the catch.

Table 4-3: Recommended fish survey measurements to determine effects in fish growth, reproduction, condition and survival.

Measurements	Expected Precision	Recommended Reporting
Length (fork or total or standard)*	+/- 1mm	Individual measurements, mean, star deviation
Total body weight (fresh)	+/- 1.0%	Individual measurements, mean, star deviation
Age	+/- 1 year (10% to be indepe confirmed)	Individual measurements, mean, star deviation
Gonad weight	+/- 1.0%	Individual measurements, mean, star deviation
Egg size	+/- 1.0%	Weight, minimum sub-sample sizes eggs
Fecundity**	+/- 1.0%	Total number of eggs per female
Weight of liver	+/- 1.0%	Individual measurements, mean, star deviation
External condition	NA	Obvious abnormalities, prevalence o lesions, tumours, parasites, etc.
Sex	NA	Male, female, immature

* If caudal fin forked, use fork length. Otherwise, use total length. In cases where fin erosion is prevalent, standard length should be used.

** Fecundity can be calculated by dividing total ovary weight by weight of individual eggs (individual egg weight can be estimated by counting the number of eggs in a sub-sample. The water displacement method is not recommended. The sub-sample should contain at least 100 eggs). The eggs should be weighed prior to preservation as the preservative will change the weight of the eggs.

4.3.2.2.4 Timing of Sampling

Timing of sampling is crucial to the success of a fish survey. For species having spawning movements that take them away from or temporarily into the effluent field, a survey during or near the spawning season would be ineffective. Thus, for spring-spawning freshwater species, a fall survey would be appropriate when residency or spawning migration is a significant issue. Early spring spawners for which it is often impossible to sample prior to the spawning season should be collected as late in the year as possible. For fall spawners, a spring or summer survey is appropriate. This may not apply to fish in which ova mature rapidly. For example, some late-spring-spawning minnows and many marine species should be sampled in early spring, rather than in fall when ova may still be immature. Some species undertake a very short period of gonadal growth and sampling schedules should optimise both residency and gonadal development concerns. Other factors to be considered in scheduling fish collections include water conditions (flow, turbidity, wave action), accessibility, and the cycle of gonadal development for the fish species.

In general, reproductive periods for molluscs and patterns of abundance are related to climate and the abundance of food supply. For most freshwater lotic or lentic habitat types, sampling is best conducted during the fall when the majority of taxa will be present and/or are large enough to be easily collected. In marine environments, sampling should be conducted in late summer or fall as populations with spring recruits have stabilized by this time.

The timing of sampling and the choice of fish species should be made according to normal operation of the mining facility to ensure that the effluent is present in the environment. Sampling when effluent has not been discharged for long periods (months) should be avoided, as mines are known to discharge at irregular rates during the year. However, the selected sampling gear, flow conditions, and effluent conditions may limit the preferred season for the survey. All sampling gear may be difficult to use during periods of high flow and it is known that some mines may have very high spring discharge rates to relieve water inventories while stream flows are high (maximum dilution power).

If historical data exists, it would be useful to examine the data and, if appropriate, conduct the survey during similar periods so that the surveys can be compared. Similarly, subsequent monitoring should be conducted during similar periods of the year to be comparable. Many factors exists (e.g., sampling season) that can confound the results if proper care is not taken to control them.

4.3.2.3 Fish Usability

The objective of the question "has there been a change in fish usability due to effluent?" is to determine whether effluent has altered fish in such a way as to limit their use by humans. Fish usability can be affected by altered appearance, altered flavour or odour (tainting) or contaminant levels which exceed consumption guidelines for human health and are statistically different from levels measured in the reference area.

Obvious visual abnormalities, such as tumors, external lesions, obvious parasites and spinal column malformations should be described and reported for each fish captured while conducting the fish survey.

Fish usability is also evaluated by measuring concentrations of mercury in tissue from fish in the exposure and the reference areas. If a mine is detecting levels of mercury greater than or equal to 0.10 μg/L as part of effluent characterization (Chapter 6), the mine is required to conduct a fish tissue analysis. Mercury is the only metal for which there is a standard Health Canada tissue consumption guideline for humans, therefore of national concern. The 1999-2000 Ontario Guide to Eating Sportfish states that Health Canada recently revised downward the tolerable daily intake for mercury, which is temporary pending the completion of a long term study. As a result, consumption restrictions for sportfish begin at levels above 0.45 μg/g total mercury. We chose the lowest guideline that was available.

In the Ontario Guide to Eating Sportfish, it is suggested that other metals (Pb, Cu, Ni, Zn, Cd, Mg, Cr, As, Se) may be found in fish tissue, but not at levels for consumption restrictions. On a site specific basis these metals may be identified as a concern if there are human health consumption guidelines from another regulatory agency (e.g., provincial or territorial), that are applicable to the region where the study is being conducted **and** if the metal for which there is a consumption guideline is present in the effluent. Local consumption and commercial fisheries should be considered to determine which edible tissues (liver, kidney, bones, flesh or even entire fish) should be used in the analysis.

A statistical difference between of the concentration of mercury in fish tissue of exposure area fish compared to reference area fish, ***and*** an exceedence of the human health consumption guideline of 0.45 µg/g wet weight would indicate an effect on fish usability. Because at some mine sites there may be reference areas that have levels of total mercury in fish tissue higher than 0.45 µg/g (e.g., northern Quebec – Schetagne et al. 1996; Schetagne and Verdon 1999) to be considered an effect in fish tissue, there must be a statistical difference between the areas and an exceedance of the guideline.

Molluscs can accumulate metals (Cd, Cu, Zn, Pb, Ni, Hg, As, Ag, and Cr). Field studies suggest that the relationship between mollusc tissue metal concentrations and ambient metal concentrations are influenced by a number of biological, physical and chemical parameters that need to be taken into account. Ultimately, the relationship is metal specific and depends on the availability of the metal from the dissolved and particulate phase (AETE 1997).

If a mine is aware of any registered complaints about fish flavour or odour (including written complaints or documented verbal complaints to the mine or government agencies concerning fish tainting), the AO should be notified. Where appropriate, a study design for a tainting evaluation may be submitted. See Section 4.10.2 and Environment Canada (1998) for guidelines on conducting a tainting evaluation.

4.3.2.4 Alternative Methods

There may be situations where the fish survey as recommended (e.g., monitoring two fish species in the exposure area and reference area) provides only minimal useful information or may not be cost effective. Examples include:

- conducting the monitoring would subject persons doing the monitoring to hazardous conditions; or
- confounding influences may lead to a scientific inability to measure the effect of effluent on fish, that the design of the EEM study could not resolve.

Under these circumstances, the mine may select an alternative option to the fish survey and/or the fish usability survey. In order to satisfy the regulatory requirements, the alternative chosen must provide information in order to determine if there is "an effect on the fish population" or "effect on fish tissue", as defined in Schedule 5.

Detailed guidance on how to conduct the alternative monitoring methods and interpret the data can be found in Chapter 12.

4.3.2.4.1 Caged Bivalves

If a fish survey cannot be carried out, either because the receiving environment cannot be sampled, insufficient finfish or shellfish are found there or there are confounding influences that could not be resolved in the study design, one of the suitable alternatives would be a caged bivalve exposure.

The caged bivalve methodology has been refined to increase its utility to characterize chemical exposure and biological effects for effluent monitoring (ASTM Manual 2000) Salazar and Salazar 1997a,b; Andrews and Parker 1999). Caged bivalves are a good alternative to the fish survey because the method combines the elements of experimental control from laboratory bioassays with field conditions. The methodology now has well known and well established endpoints (growth and bioaccumulation) and has been used in Canada and the United States in a cost-effective manner (Courtenay et al. 1998).

In many cases, caged bivalves will be able to circumvent problems with confounding discharges and historical damage such as substrate alteration. In situations where a caged bivalve exposure cannot be carried out because the receiving environment is inappropriate (e.g., strong currents, fluctuating salinities, confounding effluents, historical damage (e.g., substrate alteration), an on-site bioassay/mesocosm may be a suitable alternative.

Detailed guidance on how to conduct the caged bivalve study and interpret the data can be found in Section 12.2.

4.3.2.4.2 Mesocosms or On Site Bioassays

Mesocosms are field based artificial stream systems which can link field monitoring with sublethal toxicity testing by simulating the natural environment while allowing for variable control, such as sample sizes, sex ratios and effluent exposure (Culp and Podemski 1995; Culp et al. 1996; Dubé et al. 1998). Mesocosms attempt to recreate natural ecosystems and examine multi-species interactions. On-site bioassays refer to artificial stream studies where some environmental realism is maintained (ambient temperature, water chemistry, photoperiod) but multi-species interactions at a community or population level are not examined. Pilot studies have shown that standard EEM endpoints (juvenile growth and size changes in adult organs) as well as physiological endpoints (plasma and in vitro steroid hormone production) can be measured after exposure to effluents over a 30 to 60 day period (Dubé et al. 1997).

Detailed guidance on how to conduct the caged bivalve study and interpret the data can be found in Section 12.1.

4.3.2.4.3 Expanded Sublethal Toxicity Testing

When it is not possible to sample fish directly, it may be possible to estimate the potential of the effluent to affect fish using fish sublethal toxicity tests at environmentally relevant concentrations. It should be noted, however, that the validity of extrapolating toxicity test results to actual field effects will be assessed using the results of the EEM program, when they are available. Therefore, it remains to be determined how successfully site specific fish sublethal toxicity test results can be extrapolated to the field. See Chapter 7 for further guidance.

4.3.3 Marine Discharges

At the current time, there are few metal mines that discharge to the marine and estuarine environment. Detailed guidance is difficult to produce given the paucity of experience, especially given the poor performance of marine fish surveys during the pulp and paper EEM studies (Environment Canada 1997; ;Courtenay et al. 2002).

Some general guidance is possible, but mines are encouraged to discuss the situation with the AO early in study design. New information will come available as experience with marine mining discharges increases

When designing a marine metal mining EEM fish survey, the following should be considered:

- some marine and estuarine areas are difficult to sample and alternative approaches should be considered. There are a number of issues that can make sampling difficult, including tides and currents, high flushing rates or unsuitable habitat.
- interpretation of results is often difficult because of confounding factors.
- there can be gradients for current, temperature and salinity. In addition to the effects of these gradients on physical processes, the gradients can have consequences for physiological changes within organisms. The gradients can also affect uptake of contaminants. The physical processes themselves can affect metal availability and effluent dispersion.
- selection of reference areas can be more difficult in marine situations.
- for new developments, it will be critical to collect baseline information prior to initiating effluent discharge.
- different life stages of fish may utilize different habitats at different times of the year. If it is determined that the exposure area is predominantly used as a rearing area for early life stages, then studies should focus on whether that aspect is impaired. If the area is used primarily as a winter feeding area, then studies should focus on whether that aspect is impaired.
- species availability can be low in marine environments. In many situations, small-bodied resident fish species are available and should be investigated. In many cases, these species may be fractional spawners or live-bearers or species for which there is little background information. The absence of background information should not restrict or inhibit attempts to use these species, especially if they are abundant and resident. The

assumption inherent in an EEM program is that a fish community should be intact, with the normal abundant species present. The second priority, and underlying assumption is that a fish population which shows a growth rate, reproductive development and an age distribution indistinguishable from a reference area is unaffected.

- in situations where wild fish can not be collected, alternative approaches such as the ones already described in Section 4.3.3.4 should be considered.

4.4 Data Assessment and Interpretation

For the purposes of the EEM program, an effect is defined as a statistically significant biological difference in fish or benthic invertebrates between the exposure and reference area. Further information on Data Assessment and Interpretation can be found in Chapter 9.

Parameters to be measured, effect endpoints and statistical analyses that will be done for the alternative monitoring methods can be found in Chapter 12.

An effect on the fish population is defined as "a statistical difference between fish population measurements taken in an exposure area and reference area."

An effect on fish tissue is defined as "measurements of total mercury that exceed 0.45 μg/g wet weight in fish tissue taken in an exposure area and that are statistically different from the measurements of total mercury in fish tissue taken from a reference area."

To determine whether or not there is an effect on the fish population, statistical analyses of the data are conducted (Table 4.4a and b). The biological variables will be used to derive the following parameters:

Energy Use (Growth and Reproduction)

- growth: size (both as length and weight)-at-age, change in length frequency over time, size of young-of-year at end of growth period, whole animal wet weight and length (caged bivalves)
- reproductive investment: gonad weight (as gonad weight relative to body weight; gonad weight relative to length), fecundity (as fecundity relative to body weight; fecundity relative to length and fecundity relative to age), sex steroid levels, young-of-year survival, or % composition of young of the year

Survival

- age structure
- length frequency analysis

Energy Storage (Condition)

- condition (body weight relative to length), condition index (caged bivalves)
- liver weight (as liver weight relative to body weight and liver weight relative to length).

Sex differences in growth rate, body weight, condition factor, gonadal size and liver size are common due to differences in overall energetic requirements between male and female fish. Therefore, for all parameters, sexes should initially be treated separately when estimating variability. Immature fish should

Table 4-4a: Effects on fish.
Endpoints recommended for determining effects as designated by statistically significant differences between exposure and reference areas.

Type of Response	Effect Endpoint[1]	Statistical Procedure (dependent on endpoint)
Survival	Age	ANOVA
Energy Use	Size-at-age (body weight against age)	ANCOVA
	Relative gonad size (gonad weight against body weight)	ANCOVA
Energy Storage	Condition (body weight against length)	ANCOVA
	Relative liver size (liver weight against body weight)	ANCOVA

---

[1] For the ANCOVA analyses, the first term in parentheses is the endpoint (dependent variable, Y) that is analyzed for an effluent effect. The second term in parentheses is the covariate, X (age, weight, or length).

Table 4-4b: Endpoints to be used for supporting analyses.
These analyses are for informational purposes and significant differences between exposure and reference areas are not necessarily used to designate an effect.

Type of Response	**Supporting Response Variable**[2]	**Statistical Procedure (dependent on endpoint)**
Energy Use	Body weight (whole)	ANOVA
	Length	ANOVA
	Size-at-age (length against age)	ANCOVA
	Relative gonad size (gonad weight against length)	ANCOVA
	Relative fecundity (# of eggs/female against body weight)	ANCOVA
	Relative fecundity (# of eggs/female against length)	ANCOVA
	Relative fecundity (# of eggs/female against age)	ANCOVA
Energy Storage	Relative liver size (liver weight against length)	ANCOVA
	Relative egg size (mean egg weight against body weight)	ANCOVA
	Relative egg size (mean egg weight against age)	ANCOVA

not be mixed with mature fish for analyses. For most species, individuals with gonad size of less than one percent (1%) of body size would be considered immature although there are some notable exceptions, such as guarding species like the bullhead. The sampling period has to be adjusted to the biology (life history) of the species to avoid capturing fish prior to gonadal development for the upcoming reproductive season. However, when non-destructive sampling is to be carried out and age frequency distributions are used to assess reproductive success, then the timing of sampling is less important.

When non-destructive sampling is conducted, reproductive success will be assessed using relative age class strength, then choices for aging structures (e.g., scales) become more important but limited (Gray et al. 2002). If a non-destructive monitoring approach is conducted but a suitable aging structure can not be taken, then a length frequency distribution may be plotted as a surrogate of an age frequency distribution. Size frequency analysis can be used to examine age distributions, size-at-age data and condition factors for the fish (Gray et al. 2002). Reproductive performance will be evaluated by the relative abundance of young-of-the-year individuals. Growth can be evaluated by the size of young-of-the-year fish at the end of the growing season, and by the size of the 1+ fi Performance can also be evaluated by the relative condition (k=

---

[2] For the ANCOVA analyses, the first term in parentheses is the endpoint (dependent variable, Y) that is analyzed for an effluent effect. The second term in parentheses is the covariate, X (age, weight, or length).

$100*(wt/l^3)$) of the fish examined. A large number of areas can typically be sampled using this approach, and the mine is encouraged to sample multiple exposure and reference areas. If possible, a small number of larger fish should be sacrificed to verify ages of older individuals. If the fish species chosen is externally sexually dimorphic, it is possible to examine whether there are gender specific differences in growth rate.

4.4.1 Statistical Guidance for the Fish Survey

Descriptive statistics should be reported for all variables (arithmetic mean, median, minimum, maximum, standard deviation, standard error and sample size).

Most parameters are estimated as regressions of one variable on another using Analysis of Covariance (ANCOVA). ANCOVA combines the features of analysis of variance and regression, and can be used to compare regressions among multiple classifications (i.e. sampling areas) (Snedecor and Cochran 1980). The mean squared error (MSE; model variance, $s^2 = SD^2$ of the residuals from the regression]) of the ANCOVA model should be used to provide an estimate of variance of the fish parameters, when reasonably large samples (e.g., including 20 or more fish) are used.

Site differences in length, weight and age can be tested using an Analysis of Variance (ANOVA). The assumptions for ANOVA are that:

- the data of the two populations are normally distributed
- the variance of the two populations are equal
- the samples are independent of each others

Data should be tested for normality and homogeneity of variance before applying parametric procedure. However, most univariate normal distribution-based statistical based methods are quite robust and can support moderate violations of the assumptions. Transformation of original data will help normalize the data or homogenize the variances. Logarithmic transformation are often preferred because most biological measures are considered to operate on a log or exponential scale (Peters 1983). If the transformations are unable to produce data that meet the assumptions, then non-parametric statistics are used.

Assumptions of ANCOVA are that:

- the slopes of regression lines among areas are equal
- the independent variable (covariate) is fixed and measured without error
- the relationship has the form specified (linear regression)
- the residuals are normally and independently distributed with zero mean and a common variance.

The assumption that the independent variable is fixed is frequently violated and Draper and Smith (1981) discuss the consequences of this violation. This is likely to prove problematic only in situations where the range of the independent variable is very small. The assumption of a linear relationship can be tested for samples with multiple observations at different values of the independent variable. This may be possible for discrete variables such as age but not for continuous independent variables such as body weight (Environment Canada 1998).

It has been suggested that the range of the independent variable should be approximately the same for each 'treatment' (area). This will be difficult to assure in practice, but the violation of this should be considered when interpreting results from such cases. Some also consider a further requirement to be that the range of the independent variable be 'compact'. This serves to avoid difficulties created by outlying values of the independent variable if there is some non-linearity in the relationship.

A brief description of the application of ANCOVA for determining size-at age follows to provide a better understanding of the procedure (From Environment Canada 1998).

4.4.2 Size-at-Age

Rates of growth are commonly described by the relationship of size (as weight or length) to age. Over the entire life span of a fish, this relationship is curvilinear, with the rate of increase declining as fish approach the limit of their life span (Ricker 1975). As only adult fish are often sampled, classical growth rates cannot be calculated. Nevertheless, for the purposes of the EEM program, fish growth can be inferred from size-at-age estimates determined for each area using ANCOVA. This calculation assumes that the relationship between size and age for adult fish is approximately log-linear (log size vs. log age) (Bartlett et al. 1984).

Size-at-age may be estimated by calculating the regression relationship between body size (weight or length) and age for each sampling area (reference and exposure). It is recommended that both length and weight be used to calculate size-at-age in order to determine which provides the best fit and tightest regression. ANCOVA can then be used to compare the separate regressions between body weight or length (dependent variable; Y) and age (covariate; X). Body weight should be corrected by subtracting the gonad weight and liver weight from the body weight prior to analysis.

An assumption of ANCOVA is that the slopes of the regression lines among areas are equal. It is first necessary, therefore, to test for differences in slopes between areas (described by the interaction term of age*area from ANCOVA). In cases where the slopes are not significantly different, this indicates that the regression lines are parallel to each other i.e. that weight gain over age is similar for both areas. The next step is to proceed with the ANCOVA model and test for differences in intercepts (elevation) to investigate whether fish are proportionately heavier at any age in one area than in another. It is possible that the slopes of regressions may differ, i.e. fish from one area are gaining more weight with age (steeper slope) than fish from another area. In this situation, it is a good idea to plot the separate regression lines to obtain a better understanding of the size-at-age relationship at each area, over the entire data range (e.g. where do the lines intersect?).

If the objective of the survey was to derive an estimate of variance for the parameter, it is best obtained from the mean standard error (MSE) of the ANCOVA model. In cases where the assumption of equality of slopes was not met, and the range of ages are similar among areas, then it is not necessary to compute the MSE as the sample size has been sufficient to detect differences between areas.

4.4.3 Fecundity, Gonad Weight, Liver Weight and Condition

A similar procedure as outlined for size relative to age can be used to estimate the variability in fecundity, gonad weight, and liver weight. The same procedure is appropriate and body weight and length are used as the covariate to adjust for any differences in size in order to place the variables on a relative scale. The variability of fish condition between areas can also be estimated using the same principle. In this case, condition is defined by the relationship between body weight and body length and essentially describes how "fat" fish are at each area. A variety of indices have been used in fisheries biology to describe the condition of fish (Bolger and Connolly 1989). Calculating the ratio of one variable to another derived many of them. Examples of a few common indices are (Environment Canada 1998):

- condition factor (k) = 100 (body weight/length^3)
- gonadosomatic index (GSI) = 100 (gonad weight/body weight)
- liversomatic index (LSI) = 100 (liver weight/body weight).

In general however, investigators have become cautious about using derived variables and ratios because they assume some relationship between variables that may not exist, and because they may have unusual and undesirable statistical properties (Green 1979; Cone 1989; Jackson et al. 1990). Although these indices may be used for presentation purposes, it is preferable statistically to estimate (and analyse) the parameters based on regressions of original variables (i.e. ANCOVA) rather than ratios (Gibbons et al. 1993).

4.4.4 Mean Age

Calculation of mean age is meant as a gross reflection of the age distribution of adult fish collected from each area. Variability in mean age of fish can be estimated using ANOVA. The MSE from the model is the best estimate of variability. Site difference in length and weight can also be analysed in this fashion. It is essential that the sampling gear be consistent between the sampling areas because most sampling methods select for certain age classes.

4.4.5 Physical Abnormalities

A comparison of physical abnormalities (e.g., tumors, lesions on the body surface, obvious parasites) between exposure and reference fish can be completed using a standard Chi Square (X^2) test.

4.4.6 Order of Data Analysis

It is essential that the data initially be examined for differences in all parameters between sexes, and if immature fish are being used, between immature and mature fish. Immature fish devote proportionally more energy towards growth, and also change body size and gonadal growth relationships. The order of data analysis is largely a matter of personal preference, but our preferences will be outlined here. The steps can be divided into five stages: a) preparing the analyses, b) initial summary statistics, c) ANOVA analyses, d) ANCOVA analyses, and e) power analyses.

a) Preparing the analyses: this step includes data entry and QA/QC, including a check for data entry errors, and confirmation that all fish which are assumed to be adults are undergoing gonadal development for the next spawning season. Data that will be necessary for interpretation should also be summarized, including significant confounding factors. Other significant information will include the gear used and catch per effort calculations. This step should also include an examination of the exposure data, and determination of whether there was a possibility of fish moving between areas, or whether the reference area fish were exposed.

b) Summary statistics should be calculated by species and sex, and should include means, standard deviations, standard error, minimum and maximum values and sample sizes. The sample size collected should be compared to the objectives for the sampling program, and the variability obtained should be compared to previous studies and the values used during study design. The data should also be examined for normality and equality of variances (basic statistical assumptions).

c) ANOVA analyses should examine whether there was a statistically significant difference in length, weight, age, egg size or fecundity for either species or sex; p level; % difference; and an indication of which size was larger when differences exist.

d) ANCOVA analyses should be performed for size-at-age, condition factor (length weight regressions), gonadal weight (gonad weight-carcass weight or gonadal weight-length regressions), liver weight (liver weight-body weight and length regressions). The analyses should examine whether there were significant regressions, and if there was a significant interaction between areas. If slopes were equal, the data should be examined for a difference between areas, and which area was larger, what is the % area difference, and what was the p for slope or intercept differences. If there is an interaction, the data should be plotted to see if the data are interpretable.

e) The statistical power of the comparisons should be examined, and sample sizes defined for subsequent collections.

4.4.7 Data QA/QC and Analysis

Below, guidance is provided on QA/QC for data analysis. The importance of ensuring data quality can not be overstated. See Section 4.11 provides further guidance on QA/QC for study design, consistency of methods and measurements, and definitions of protocols and procedures..

4.4.7.1 Data Errors and Outliers

There are various types of common entry errors, including data entry errors, entering the wrong species, missing or moved decimal places, and wrong sex or stage of maturity. It is critical to examine the data for errors and outliers prior to initiating analysis of data. Entry errors, transcriptions errors and invalid data are impossible to detect in final reports.

Data that has been entered incorrectly can be easily detected using scatterplots of length versus weight, and weight versus gonad weight and liver weight can be used to look for points that are obviously different. Data entry errors are relatively easy to correct and can be re-entered. If the error can not be reconciled because of obvious errors or omissions in the original data sheet, the fish (data point) should be removed from the data set.

Errors and extreme observations inflate the variance and reduce the power of the data set. Evaluation of outliers includes consideration of the raw data, the field conditions, and the data collection process. Data which are different, but are not due to entry errors can arise for a number of reasons. For example, fish may appear sick or damaged, the fish may be an outlier for no apparent reason, or the outliers may represent an important phenomenon which is part of the response to the stressors under study.

In the first case, there can be a small number of fish that are obviously sick or were damaged (in a manner unrelated to the stressors under consideration) and should not be considered part of the data set for interpretation. These usually appear as single points that are separate from the main data set. Examples of these include fish which are missing their tail due to predation wounds, fish that have a jaw deformity or injury which has affected their feeding, or fish which are blinded through injury and are thinner than other fish. In these cases, the fish should be removed from the comparison.

If there is no obvious reason for the presence of rare outliers, the analysis should be conducted with and without the suspect observation, to determine how much influence it has on the conclusions. If it has an impact on whether a relationship is significant or not, statistics textbooks should be consulted for advice on how to evaluate whether the measurement can be removed.

In the third case, there can be several fish that are obviously different, but may be part of the relationship is being examined. In other cases, fish can have a delay in sexual maturity associated with environmental stressors. In this case, several fish would appear as outliers. As mentioned above, the analyses should be conducted both with the outliers (to see if there are differences between sites), and without the outliers (to see if the fish showing gonadal development are showing normal levels of gonadal development).

There may be cases when some fish within a population are different, for example, in situations where some fish may skip a year of spawning. If one is evaluating impacts on spawning, the analysis should consider the potential impacts on spawners, and non-spawners independently. Individuals that skip reproductive seasons can usually be identified as negative outliers from ANCOVA. In other words, plots of residuals from ANCOVA will be skewed left, and will not be normally distributed. Residual diagnostic procedures outlined in Environment Canada (1995) will usually identify if some individuals are skipping reproductive years. These individuals should be excluded from analyses of reproductive, and possibly all, variables. The reductions in variance achieved will usually compensate for any loss of power from reduced sample sizes. If females skipping reproductive years are excluded, that exclusion should be made objectively (Environment Canada 1997). Also, the frequency of such individuals in reference versus exposed areas should be provided, in case skipping reproductive years is related to exposure. It is much more difficult to identify males that might skip reproductive years, if in fact that ever occurs.

4.4.7.2 Validity of Statistical Assumptions

The most important statistical assumptions that should be addressed are related to a) independence of data, b) normality and c) equality of variances.

4.4.7.2.1 Independence (pseudoreplication)

When designing experiments, it is desirable to ensure that replicates are randomly allocated to different treatment levels such that the responses of each replicate is independent of other replicates. This element of randomness provides some assurance that observed differences in responses among treatments results from treatment-effects and not for some other reason.

Randomly allocating replicates to different treatment levels is a relatively easy procedure when conducting manipulative experiments (e.g., controlled laboratory tests), but is less obvious for observational field studies. Observational studies, such as environmental impact studies (e.g., single stressor EEM studies) or environmental assessements (i.e., multiple stressors), test hypotheses about the presence and magnitude of effects. However, the strength of inferences from these type of experiments is limited for two reasons (from Paine 1998):

- The stressor (e.g., mine outfall, hydroelectric dam) cannot be replicated
- Stressors cannot be applied randomly to replicates

What this means is that the stressor or treatment is always partly or wholly confounded with space or time and that the observed effects may or may not be caused by the stressor of interest. For example, when investigating whether effluent from an industrial plant is having an effect on downstream fish populations, it is not possible to replicate the treatment of effluent exposure (i.e., there is only one plant and outfall), nor randomly assign fish populations to the different treatment levels (reference *vs* exposed). As such, when significant differences are observed between reference and exposed fish populations, it is only possible to conclude there are differences between these two populations, not that the differences were caused by effluent exposure. Interpreting significant differences as treatment-effects when either treatment is not replicated or replicates are not independent is referred to as pseudoreplication (Hurlbert 1984).

It is critical that observations be confirmed, through replication over time, and that some effort be expended to confirm that the stressors of interest are involved in the responses before attributing cause to any specific stressor.

4.4.7.2.2 Normality and homogeneity of variances

Data should be tested for normality and homogeneity of variance before applying parametric analyses. Most of the univariate statistical methods are robust under moderate violations of assumptions, with some exceptions, including analyses with small and unequal samples sizes or one-tailed tests (see Environment Canada 1995 for a more detailed description). Transformations of the data can reduce some violations, and an examination of the relationship between the means and variances can help to identify the most appropriate transformation (see Green 1979).

4.5 Subsequent EEM Studies

4.5.1 Periodic Monitoring

Periodic Monitoring can be carried out at three different levels: confirmation, surveillance and minimal monitoring. The major differences among the three phases are the sampling frequency and the triggers.

For the EEM program, an effect on the fish population is defined as "a statistical difference between fish population measurements taken in an exposure area and reference area." An effect on fish tissue is defined as "measurements of total mercury that exceed 0.45 μg/g wet weight in fish tissue taken in an exposure area and that are statistically different from the measurements of total mercury in fish tissue taken in a reference area. With respect to this definition, a mine will proceed with Periodic Monitoring to:

- ensure effects are replicable, consistent and mine related (confirmation) – interpretative report submitted within 24 months of previous interpretative report;
- ensure there are no effects (after one survey) or to ensure that confirmed effects for which the magnitude, extent and cause are known are not changing (surveillance) – interpretative report submitted within 36 months of previous interpretative report;
- ensure there are no new effects, after no effects were found in fish and the benthic invertebrate community following two consecutive surveys (minimal) – interpretative report submitted within 72 months of previous interpretative report.

Monitoring requirements for Periodic Monitoring are the same as those for the first EEM study population measurements and fish usability. The frequency of biological monitoring is dependent on the results of the previous biological monitoring.

At any time, if effects are seen in either the fish survey or benthic invertebrate community survey, the mine may propose to proceed directly to Focussed Monitoring or Investigation of Cause.

4.5.1.1 Confirmation Monitoring

If effects are found in the previous monitoring, then the mine is required to confirm these effects. The purpose of the confirmation phase is to confirm the presence of effects, when effects have been found in the fish, fish tissue and the benthic invertebrate community, in the preceding monitoring phase. Confirmation monitoring is completed, including the submission of the Interpretative Report, no less than 24 months from the submission of the preceding Interpretative Report *(Schedule 5, s. 22(2)(a))*. Sampling should be done in the same time frame (e.g., the same sampling season) as previous monitoring. Mines may redesign the study before confirmation monitoring if it was questionable whether the effect seen was mine related. A mine can move directly to Focussed Monitoring or Investigation of Cause if an effect is found and it has not been confirmed. Once an effect is confirmed, the effect is considered mine related and the mine will proceed to Focussed Monitoring or Investigation of Cause.

4.5.1.2 Surveillance Monitoring

Surveillance monitoring is conducted when:

- no effects in the fish, fish tissue and the benthic invertebrate community have been found; or,
- an effect has been found in one or two of the fish, fish tissue or benthic invertebrate community (but not all 3) in the previous one study; or,
- to continue to monitor effects when an effect has been found and the magnitude, geographic extent and cause of the effect are known.

The surveillance phase of Periodic Monitoring is completed, including the submission of the Interpretative Report, no less than 36 months from the submission of the preceding Interpretative Report *(Schedule 5, s. 22(1)*.. Sampling should be done in the same time frame (e.g., same sampling season) as previous monitoring.

4.5.1.3 Minimal Monitoring

If after two consecutive EEM studies, no effects have been seen in the fish, fish tissue **and** benthic invertebrate community, the mine may conduct minimal monitoring. The minimal monitoring phase of Periodic Monitoring is completed, including the submission of the Interpretative Report, no less than 72 months from the submission of the preceding Interpretative Report *(Schedule 5, s. 22(2)(b))*.. If there are permanent changes in the mine operations, including permanent changes to the locations of the planned final points of discharge, a permanent physical change to the study area, or a permanent significant increase in loadings of items 1 to 8 of Schedule 3 of the MMER, confirmation monitoring of both fish and benthic invertebrates is recommended no less than 24 months from the date of the change.

4.5.2 Focussed Monitoring

Focussed Monitoring is conducted to determine the magnitude and the geographic extent of an effect. Due to the site specificity of the program, there is no generic guidance for Focussed Monitoring except the general criteria for good scientific practice already discussed.

Focussed Monitoring involves increasing sampling effort in terms of number of areas (geographic extent) or number of stations (magnitude) progressively further from the final point of discharge until no effect is found, to provide more information specifically to answer the question "is the magnitude and geographic extent known?".

Studies related to determining the magnitude should focus on determining the size of the change further from the discharge, in the exposure area. It may be that the samples collected during Periodic Monitoring may be sufficient for determining the magnitude. However, it might be important in some circumstances to collect additional information.

In the case of growth, it may be helpful to collect information on other age classes. For example, it may be important to determine whether there are changes in growth of early life stages. This will assist in determining the magnitude of the effect. In terms of a change in gonad size, additional work related to magnitude could be focussed on determining whether the change occurs at other times of the year (for multi-spawners). In both examples, studies related to magnitude could focus on asking whether the changes are present in other species in the same area.

Focussed monitoring is required to be completed within 24 months of the submission of the previous interpretative report *(Schedule 5, s. 22(2)(c))*.

4.5.3 Investigation of Cause

Studies may progress to Investigation of Cause prior to confirmation of the effect or determining the magnitude and extent. However, the mine can determine the magnitude and the geographic extent of the effect prior to initiating Investigation of Cause.

The Investigation of Cause is required to be completed within 24 months of the submission of the previous interpretative report *(Schedule 5, s. 22(2)(c))*.

The following are possible tools that can be used during Investigation of Cause.

4.5.3.1 Ecoepidemiological Criteria for Attributing Cause

There are a number of papers that discuss the principles of using ecoepidemiological techniques to identify causes of effects (Fox 1991; McMaster et al. 1996). These types of studies are common with field studies, but the results are often indirect attribution of causes.

Ecoepidemiological studies determine lines of evidence for defining cause and effect. The following should be considered (modified from Fox 1991):

Category	Issues
• Consistency	Are the responses consistent? (with different times, places, researchers, species, contaminants)
• Predictive Performance	Does the hypothesis predict a previous unknown fact or consequence?
• Probability	Statistical strength of conclusion; is there a plausible mechanism for the relationship?
• Time Order	Is there consistency in the timing of the proposed cause and effect?
• Strength of Association	What is the relative risk of increased exposure?
• Specificity	What is the precision of the relationship; does the exposure lead only to that effect and be caused by only that exposure?
• Coherence	Is there consistency between cause and effect relationship and known facts? The cause-effect history should not conflict with facts and theory.

4.5.3.2 Identification of Specific Stressors Primarily Responsible for Changes

Once effects have been confirmed and the magnitude and extent of the effects are known, it is possible to eliminate or prioritize stressors based on the geographic distribution of responses. The studies may include using:

- gradient designs. These can be used when there is geographic distance between stressors.
- caged fish/non mobile fish or invertebrates. These can isolate factors.
- on site bioassays/mesocosms. These are helpful for identifying specific effluent contributing factors.
- sediment quality triad. This has been used to provide confirmation that changes are due to site specific situations.

4.5.3.3 Identification of Specific Chemicals Associated With Changes

In order to truly identify components responsible for changes, the contaminant of concern should be removed and subsequently recovery must be found in the laboratory. The stressor is then reintroduced to attempt to recreate the impact in the laboratory. Some of the methods that can be used include:

- selective operation of the discharge or in-plant processes; or
- the use of toxicity identification and evaluation procedures (TIEs).

4.6 Methods of Sampling

The same sampling methods can be used for population and community surveys. The difference lies with the selectivity of the fishing gear. During a community survey, the gear should be as non-selective and non-destructive as possible. For the population survey that focuses on one or two species, the gear will be more selective. For example, trap netting may be preferred during a community survey while a one-size mesh gill net of the proper size could be appropriate for a population survey.

Sampling methodology should be designed site-specifically, and capture gear and effort should be focussed on methods shown to be successful. Standardized sampling is a priority. Therefore, in situations where sampled species are the same as for a previous survey, and the sampling techniques used previously were sufficient to capture the target number of each species, these same sampling techniques should be retained unless there are good reasons for change, such as unacceptable by-catch.

4.6.1 Gillnets

Gillnets are generally appropriate in rivers, lakes and deep reservoirs. During a fish survey, the correct mesh size should be selected to catch fish of the appropriate species and size, and to limit unwanted by-catch. Gillnets should be checked frequently to free unwanted fish species. Fishing may be more successful at night, depending on the fish species sought.

Gillnets are not recommended where there is unacceptable high by-catch. There may be major logistical problems with this technique because tidal or river currents can cause twisting of the nets. If gillnets are chosen as a technique, they can be sunk to fish at particular depths or floated to sample the surface layers only. Since salinity and temperature typically stratify the pelagic zone in marine environments, the fish species below and above the halocline are usually quite different. If the gillnet is several meters deep, the net will sample both below and above the halocline and the exact habitat where a particular fish species was caught may be difficult to determine.

4.6.2 By-Catch

It may be possible to obtain samples using the by-catch of commercial, research or other fisheries operations in either marine or freshwater situations. The investigator is responsible for ensuring and documenting that sampling procedures and conditions are met (QA/QC), and that fish are exposed, within the exposure area. Capture techniques have to be standardized between areas.

4.6.3 Electrofishing

This method is useful in shallow rivers and lakes (2 m and less) and may utilize a backpack or boat-mounted shocker. Unwanted by-catch can be released. Procedures that may enhance success include operating at dusk or night, passing over the same area at least three times, and using intermittent pulses of current (since a continuous field may actually chase fish away). It is important to ensure that the captured fish are representative of the population. Daylight sampling of shallow nearshore areas does not always produce representative samples. Because of its non-specificity, this technique can be used for both population and community surveys. Proper adjustment of the current regulator will keep mortality to a minimum and fish can be kept alive and released if necessary.

4.6.4 Trawling

Trawling may be appropriate for collecting bottom fish and some bivalves in coastal areas with a fairly flat seabed. On coastlines characterised by steep bathymetry (e.g. fjords), diffusers often discharge at depths greater than 20m and so introduce effluent into a steep underwater landscape. Trawling in these areas can be difficult because of the steep angle of the bottom. In these circumstances keeping an otter trawl open using the spreading forces of the otter boards may be difficult. Thus, beam trawls that use wood or metal frames to keep the net open may be preferable.

Underwater debris such as logs and rocky outcrops can also complicate trawling in fjords. Heavy mesh trawls resistant to tearing from snags should be used. The mesh size of the trawl is dependent on the species and size of fish targeted in the survey. Because changes in depth and substrate may result in changes in species composition of fish, it is important to ensure that the beginning and ending of trawls are precisely located using differential GPS. This should ensure that the same habitat is sampled in repetitive surveys. Other important variables such as time of start and stop of the trawl, depth, and substrate should also be noted.

4.6.5 Beach Seines, Fyke Nets, Trap Nets, Hoop Nets and Fish Traps

When effluent is discharged into rivers, shallow shorelines or intertidal habitat, seining may be considered. Permanent markers such as small cairns should be set up to mark the seining locations. Depending on the steepness of the shore, a small boat or persons wearing chest waders may be used to pull the net off the shore. Beach seines are routinely used for sampling small fish in rivers and estuaries and on beaches. They are most effective at night. Trap nets and seine nets are generally less selective than gillnets but they are non-lethal and should be preferred. Trap nets also tend to be very effective in shallow waters (1-2 m). On the Atlantic coast, fyke nets set in the intertidal zone are also an option. Baited traps are also used to catch mummichogs in Atlantic estuaries.

4.6.6 Hook and Line and Hand Collections

To minimize the impact of fishing gear on habitats, hand collection may be suitable for relatively sedentary species. This type of sampling can be very selective so that only the target species is sampled. For example, intertidal blennies and cottids can be collected at low tide by searching under rocks and algae. Bivalves can also be collected by hand. SCUBA is a technique that allows sampling to 30 m. Other selective methods, such as a string of baited hooks (longlining) for bottom fish such as sole, should also be considered. Angling should also be considered since it can be selective, non-lethal and remarkably effective at sampling the targeted species. As with the other techniques, in order to guarantee the exact habitat can be revisited in subsequent phases, precise recording of positioning is essential.

4.6.7 Purse Seines and Midwater Trawls

Purse seining and midwater trawls may be used to sample species that live primarily in the pelagic zone. In order to effectively sample fish from depths where effluent may be influencing fish living near bottom or in deeper waters, a fairly deep seine (e.g. 10 m deep) may be used. A large vessel such as a commercial salmon or herring seiner may have to be deployed. Midwater trawls have an advantage over seines in that they allow more precise layers of the water column to be sampled. In addition, the depth that the net is being towed at can be recorded and/or monitored electronically, so that adjustments can be made as needed.

4.6.8 Remote Sensing

Video or still camera recording of fish abundance near outfalls can be monitored using cameras mounted on remotely operated vehicles. This technique may be particularly effective in rocky and steep areas where use of fishing gear may be difficult. Camera surveys may also be useful in reconnaissance surveys of bottom conditions before trawls or traps are deployed for fishing.

4.6.9 Determination of Sampling Effort

Examples of fishing methods that have been successful in collecting fish in a timely manner are provided below. They are examples to guide consultants in the development and implementation of their study design, and to indicate when it may be advisable to consult with the AO.

Data from two Ontario lakes indicated that 40 individuals of any of 6 warm waters fish species were collected in from 1 to 6 sets of 24 hours duration. The equipment included a 6 x 6-foot trapnet. Details of the process are presented in the Ontario Ministry of Natural Resources Fisheries Assessment Unit Newsletter (FAU Update Issue 94-1, 1994).

During the Assessment of the Abundance of Cold Waters Ontario Fish Communities Program, the fishing effort recommended to collect 40 lake trout in 7 lakes varied from approximately 12 to 120 hours. The equipment included a 46 m gillnet gang with three 15.2 m panels. Details of the process are in the Ontario Ministry of Natural Resources Fisheries Assessment Unit Newsletter (FAU Update Issue 94-2, 1994). Mesh size should be consistent and selected according to the target species.

Experience has shown that gill nets made up of 4 panels of 50 meters of net, set 24 hours a day for 5 days (equivalent to 24,000 meter-hours of effort) in freshwater systems, should allow for 20 fish of each sex to be collected. This is contingent on correct deployment of the panels and shifting of the panels to cover areas inhabited by the fish. Mesh size should be consistent and selected according to the target species.

Alternatively, a good strategy would be to initially set a minimal amount of net to decrease by-catch (< 400 m). If fishing is selective enough, and the amount of by-catch acceptable, up to 2 km of a single mesh size has been necessary in small unproductive rivers.

In marine situations, experience has shown that 48 hours of beam trawling, long-lining using a variety of hook sizes or other methods, including traps, (alone or in combination) should provide the 20 fish of each sex.

Consultation with users of electrofishing (large rivers using boat-mounted apparatus) technique indicates that all fish can be obtained in one day. Procedures enhancing success include operating at dusk or night, passing over the same area at least three times, and using intermittent pulses of current (since a continuous field may actually chase fish away). In small streams, lakes and rivers more time will often be necessary for sampling because of the difficulties encountered in moving through these environments.

Although the above techniques and gear may apply to a variety of species, these examples are not all inclusive since each site is unique and are provided as suggested effort only. Local expertise can provide further advice. The previous examples are adequate guidelines toward catching a minimum of 20 fish per sex, per species and per area. Catch per unit effort (CPUE) should be reported in the interpretative report.

4.7 Sampling and Measurement Precision

For each fish collected during the fish survey, the recommended measurements are of the following variables:

- length;
- total body weight;
- weight of liver;
- age;
- sex;
- weight of gonads;
- egg size/fecundity estimate for mature females; and
- external condition.

The precision and the measurement that should be reported are detailed in Table 4.3.

For the length of the fish, depending on the selection of the species, different lengths may be taken. If the fish has a fork-like tail, fork length should be recorded. If not, total length should be used. In the some cases, standard length may also be measured. Note on the data sheets which length was recorded.

In addition, visual qualitative observations should be made of the external condition of each fish, noting obvious abnormalities, prevalence of lesions, tumours and parasites, etc. in order to determine the condition of the fish. A sample checklist for the physical examination of fishes is presented in Figure 4.2 and Table 4.5.

A list of appropriate ageing structures for a variety of fish species has been provided in Table 4.6. Formal age validation is recommended as part of the fish survey and methods of ageing should be consistent at each sampling area and among monitoring phases and appropriate QA/QC procedures followed (e.g. independent confirmation). It is recommended that all ageing structures be archived for future reference.

Small-bodied fish usually have a shorter life span than larger fish, but there is generally a greater success with traditional age structures such as scales (Table 4.6). With small-bodied fish, additional structures (such as opercula and otoliths) should also be preserved in case the scales should prove inadequate or verification is necessary.

Figure 4-2: Features of composite fish.

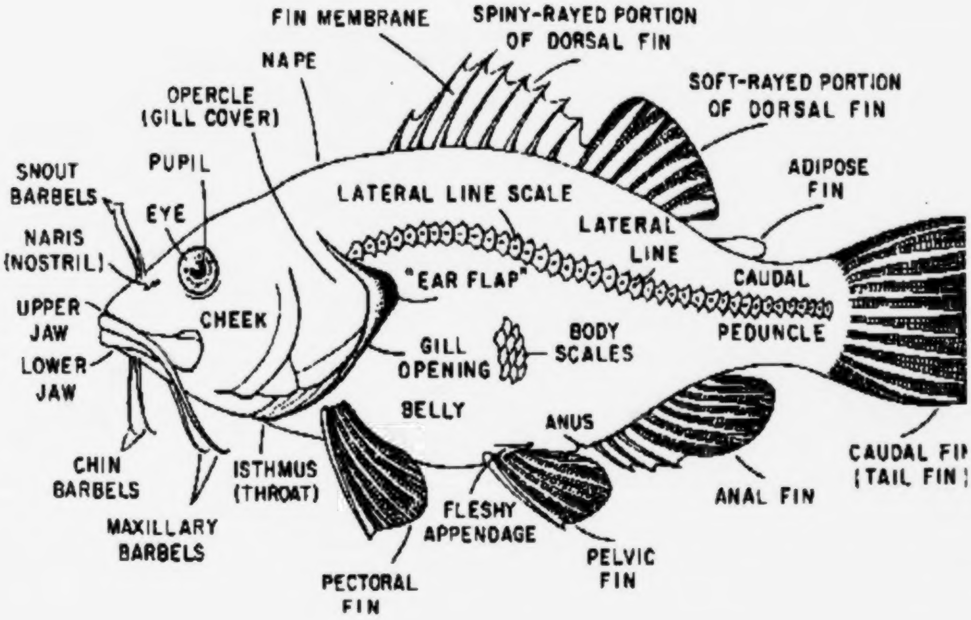

FIN MEMBRANE
SPINY-RAYED PORTION OF DORSAL FIN
NAPE
SOFT-RAYED PORTION OF DORSAL FIN
OPERCLE (GILL COVER)
PUPIL
SNOUT BARBELS
EYE
LATERAL LINE SCALE
ADIPOSE FIN
NARIS (NOSTRIL)
LATERAL LINE
"EAR FLAP"
CAUDAL PEDUNCLE
UPPER JAW
CHEEK
LOWER JAW
GILL OPENING
BODY SCALES
BELLY
ANUS
CHIN BARBELS
ISTHMUS (THROAT)
CAUDAL FI
(TAIL FIN)
ANAL FIN
MAXILLARY BARBELS
FLESHY APPENDAGE
PELVIC FIN
PECTORAL FIN

Table 4-5: Sample Data Checklist for Fish Collections including Physical Abnormalities

This represents a list of the information that should be collected and included on field data sheets for fish collections. Data sheets should include:

Date:
Location:
Collection method:
Recorder:
Species:

Fish number:
Sex:

Length – Fork, Total or Standard (mm):
Total Weight (g):

Gonadal Condition:
Gonad Weight (g):

Liver Weight (g):

Tissue Samples:
- Whole Body:
- Liver:
- Kidney:
- Digestive Tract:

4.8 External Appearance

It is recommended that a rough illustration of the selected fish species be incorporated into the data collection sheet at each mine location for the recording of abnormalities in the external appearance. If significant differences in external appearance exist, these will become evident and a clear picture of the abnormality will be recorded. This information can then be used by others at a later date if significant differences exist between reference and exposure areas.

Examination of the external appearance should include body form, body surface, fins, eyes, lesions etc. Figure 4.2 gives a representation of what parts of the fish should be examined. An area on the data sheet should also be included for other significant observations. Photographs can be a useful tool to document any obvious abnormalities on fish caught.

Table 4-6: Anatomical structures commonly used for age determination (from Environment Canada, 1998)

Family	Common Name	Structure
Acipenseridae	Sturgeons	Otoliths, pectoral fin ray
Anguillidae	Freshwater eels	Otolith
Atherinidae	Silversides	Otolith
Batrachoididae	Toadfishes	Otolith
Bothidae	Lefteye flounders	Otolith, scale
Caranigidae	Jacks	Otolith
Catostomidae	Suckers	Fin ray, opercular bone
Centrarchidae	Sunfishes	Scale
	Bass	otoliths, 1st three dorsal spines
Cichlidae	Cichlids	Scale
Clupeidae	Herrings	Otolith
Coregoninae	Whitefish	Otoliths, first four marginal pectoral fin rays, scales
Cyprinidae	Minnows and carps	Otolith, scale
Cyprinodontidae	Killifishes	Scale
Esocidae	Pikes	Scale, cleithra
Gadidae	Codfishes	Otolith
Haemulidae	Grunts	Otolith
Ictaluridae	Bullhead catfishes	Pectoral spine
Lophiidae	Goosefishes	Vertebrae, ray
Malacanthidae	Tilefishes	Otolith
Mugilidae	Mullets	Scale
Percichthyidae	Temperate basses	Scale
Percidae	Walleye, perch	Otoliths, first three dorsal spines, opercular bone, scale
Pleuronectidae	Righteye flounders	Otolith, scale
Rajidae	Skates	Vertebral centrum
Salmonidae	Lake trout	Scale, otolith, subopercula, 1st four marginal pectoral fin rays
Sciaenidae	Drums	Scale
Scombridae	Mackerel	Otolith, fin ray
Serranidae	Sea basses	Scale
Sparidae	Porgies	Scale
Squalidae	Dogfish sharks	Dorsal spine

4.9 Special Considerations for Small Bodied Fish Species

Many small species have fewer, larger eggs. Gonadal estimates will be much easier closer to spawning. The timing of sampling will also be affected by residency, and the two factors have to be optimised. The entire gonad should be preserved and fecundity counts conducted with aid of a dissecting microscope. Basic information on gonadal development should be available for many species through general references (Scott 1967; Scott and Crossman 1973; Fritz et al. 1975; Roberts 1988; Nelson and Paetz 1992; Jenkins and Burkhead 1993; Coad 1995; Leblanc and Couillard 1995; Leblanc et al. 1998).

Species identification of small-bodied fish should be verified, especially for cyprinids that can appear very similar without careful examination (Scott 1967; Scott and

Crossman 1973; Roberts 1988; Nelson and Paetz 1992; Jenkins and Burkhead 1993; Coad 1995). The smaller organ size of these fish might require a more sensitive balance. Dissecting microscopes may be necessary for removing the organs properly and avoiding extraneous tissue or moisture that could affect results. Dissection on fresh fish is highly recommended over frozen or preserved specimens. Differentiation and separation of tissues is easier when the tissue is fresh, and separation of the liver and gonads from intestinal tissue is less time consuming. Dissection of frozen specimens of small fish can be difficult and lead to errors in organ measurements. However, preservation in a formalin solution may give adequate results.

Live-bearers require special attention regarding measurement of reproductive variables, but success has been seen with detecting effects using live-bearing species in exposures to Swedish pulp mills (Vetemaa et al. 1997). The estimate of fecundity requires that the gonads be preserved and the number of live and dead embryos be counted. Proper sampling requires some preliminary data on spawning time and gonadal development so that sampling procedures can be optimised. These data may not be available for periodic monitoring of live-bearing fish, but should be evaluated on a case-by-case basis.

4.10 Fish Usability / Exposure Method

4.10.1 Fish Tissue Analysis

Fish usability is also evaluated by measuring concentrations of mercury in tissue from fish in the exposure and the reference areas. If a mine is detecting levels of mercury greater than or equal to 0.05 μg/L as part of effluent characterization (Chapter 6), the mine will do a fish tissue analysis. Mercury is the only metal for which there is a standard Health Canada tissue consumption guideline for humans, therefore of national concern. The 1999-2000 Ontario Guide to Eating Sportfish states that Health Canada recently revised downward the tolerable daily intake for mercury, which is temporary pending the completion of a long term study. As a result, consumption restrictions for sportfish begin at levels above 0.45 μg/g total mercury.

In the Ontario Guide to Eating Sportfish, it is suggested that other metals (Pb, Cu, Ni, Zn, Cd, Mg, Cr, As, Se) may be found in fish tissue, but not at levels for consumption restrictions. On a site-specific basis these metals may be identified as a concern if there are human health consumption guidelines from another regulatory agency (e.g., provincial or territorial), that are applicable to the region where the study is being conducted **and** if the metal for which there is a consumption guideline is present in the effluent. It is recommended that other metals in fish tissue be analyzed where there area site specific concerns.

4.10.1.1 Tissue Sample Collection and Preparation

Tissue analyses should be conducted on 8 samples (to achieve 95% power) of a single species from the exposure area and the reference area. If possible, the samples should be of **one** sex and age class. If this is not possible, then the sex of each fish making up the sample should be reported. If fish are not of the same age class, the age classes of the fish should be consistent between the sampling areas.

Each sample should consist of a minimum of 50 g fresh weight of tissue to provide sufficient material. The 50 g sample may be from one fish, or a number of fish to achieve the 50 g sample. Fish should be used independently in a sample and not mixed between samples. Each sample should be clearly labeled and forwarded to the analytical laboratory for determination of specified data. Samples are to be homogenized and a subsample taken for metals analysis.

The species selected for tissue analyses should be, if present, sport, subsistence and/or commercial species, including molluscs and crustaceans, where relevant. The fish species used for the tissue analysis may or may not be the same as the species used in the fish survey. On a site-specific basis, the tissue used for the analysis should be chosen based on the portion of the fish constituting the edible portion locally consumed, including the muscle, liver, eggs, hepatopancreas (crustaceans), bone or any other relevant portion.

Tissues collected for analysis should be handled to avoid contamination from sources such as survey boat fuel. For finfish, edible portions (e.g., fillet, eggs, liver) should be kept frozen until analysis.

For molluscs, whole soft body parts should be collected, and it may be necessary to composite tissues from more than ten specimens to create an adequate sample weight. For lobster or crab, edible tissue (e.g., muscle, eggs, hepatopancreas) should be collected and samples should be handled as outlined above for fish tissues.

4.10.1.2 Supporting analyses : Lipid and Moisture %

Monomethylmercury (MeHg) comprises almost all (95% or greater) of the total mercury (inorganic mercury and methylmercury) found in muscle tissue of fish regardless of the composition of diet sources and exposure water (Bloom 1992). Because of its strong affinity for sulfhydryl groups of proteins, the relative ease with which it passes through the digestive wall, and slower depuration rate relative to inorganic mercury, MeHg is accumulated and retained in biological tissues (Clarkson 1994; Saouter et al. 1993). MeHg is not highly lipohilic, as often reported, with moderate log Kow values up to 2.5 found experimentally (Halbach 1985; Major et al. 1991).

Lipid concentration has been used to normalize tissue residues among species or within species between seasons, as well as being a key variable in modeling bioaccumulation. Lipid extraction methods by Randall et al. (1991) and the chloroform-methanol extraction method developed by Folch et al. (1957) and modified by Bligh and Dyer (1959) are recommended. Lipid analysis should only be completed when the contaminant being tested is known to be lipophilic.

Percent lipid and percent moisture determinations should be provided for every sample submitted for total metal analysis. Also percent lipid values should be reported for the replicates analyzed in the same batch with the submitted sample. The percent lipid precision for the replicate samples should be ±30% for tissues containing more than 2% and ±60% for tissues with less than 2% lipid. The method for the lipid determinations would be reported and the solvents used clearly specified.

4.10.2 Tainting Evaluation

The definition of tainted fish used in this document is based on the ISO 5492: 1992 Sensory Analysis Vocabulary and is defined as "fish that contains an abnormal odour or flavour". The context of the definition is that the odour or flavour is introduced into the fish from outside sources and does not include any natural by-product from the deterioration in quality due to ageing of fish in storage, decomposition of protein, fat, or other components, or due to microbial contamination normally found in fish.

Because the decision to investigate tainting associated with a mine is expected to occur on a site-specific basis, the local knowledge and expertise of regional authorities, such as Department of Fisheries and Oceans fisheries biologists or provincial ministry district biologists, is critical for ensuring that there is a sound basis for triggering tainting evaluations. The need for tainting studies should be based on a range of pertinent information including historical information, records of complaints, supporting information associated with a record of complaint, any previous tainting studies that were undertaken, and any subsequent environmental alterations or changes to mine operations that might be relevant.

Thus it is incumbent on the local regulatory authorities, taking into account the mine's operating requirements and the interests of all stakeholders, to assess whether a tainting study should be recommended as part of an EEM phase.

In cases where the responsible parties or the mine ownership do not have immediate knowledge of all prior information from which to assess whether tainting evaluations are recommended, some effort might be needed to substantiate and/or validate existing information. In a similar manner, it might be necessary to confirm that the lack of complaints or documentation pertaining to a tainting problem is not a consequence of the abandonment of previously utilized fisheries resources by local users. A lack of utilization of fisheries resources should be due to tainting and not simply due to the perception that because a mine is in the vicinity it is not worth fishing there. At the same time, a lack of complaints should be due to a lack of tainting and not due to the fact that local users have not utilized the resource for many years because the fish were tainted at some time in the past. In most cases, such situations will be readily apparent, but there might be instances where further substantiation or validation is recommended.

When further substantiation or validation is recommended, a number of considerations may be necessary. Of primary importance is whether it is possible to assume that fish tainting is attributable to the discharge of effluent. Therefore, the site of the mine discharge, the location of natural and other anthropogenic discharges, physical characteristics (e.g. dilution capacity) of the receiving water body, and the location and utilization of fisheries resources near the mine should all be considered. Confirmation of the location and utilization of fisheries resources may require further investigation. For example, local libraries often have published histories of the region which could assist with understanding the availability and utilization of fisheries resources. Regulatory agency records could also contain relevant information. In cases where there is a history of complaints, the source and validity of such complaints should be reviewed. Creel records and/or surveys of local fishers can provide some guidance as to the basis of complaints. However, care should be taken to ensure that these records are unbiased and reproducible.

In cases where tainting evaluations have been undertaken in the past, the results and methods employed should be carefully reviewed.

See Environment Canada (1998) for the recommended protocol for a tainting evaluation.

4.11 Quality Assurance and Quality Control

Quality assurance (QA) encompasses a wide range of internal and external management and technical practices designed to ensure data of known quality commensurate with the intended use of the data.

Quality control (QC) is an internal aspect of quality assurance. It includes the techniques used to measure and assess data quality and the remedial actions to be taken when data quality objectives (DQOs) are not realised. Within the context of a particular study, assurance of adequate data quality is only possible when DQOs have been defined. Users of the data should play a lead role in defining DQOs for a study and in ascertaining whether laboratory quality control limits are consistent with these objectives.

The concepts and basic elements of a good measurement system have also been discussed elsewhere (e.g. OECD 1981, 1989; ASTM 1988; Environment Canada 1991). Examples of checklists are provided for evaluation of QA/QC for aquatic biology studies (Table 4.7)

Quality assurance functions, personnel responsible for each QA function and corrective actions when performance limits are exceeded should be identified in the quality management plan.

Quality control activities define the boundaries of acceptable performance for the measurement system, and include the routine checks (data quality measures) that indicate whether the system is performing to specification. Data reporting generally stops and corrective actions are initiated when the system goes out of control. Range and average control charting methods have been described elsewhere (OMOE 1984; ASTM 1985, 1986; Dux 1986).

Data quality measures should be defined in the same terms as data quality objectives (DQOs), so that the two can be compared in project evaluation (QAMS 1986). Data quality objectives (DQOs) are normally derived from intended data uses (e.g. hypotheses to be tested, summary statistics involved and total uncertainty that can be tolerated). Total uncertainty includes imprecision (sampling, analytical, environmental) plus any analytical bias that may occur (Taylor 1987). Objectives can be established for each component, as well as for total uncertainty, and should be incorporated into the QA project plans. The various components of imprecision can be estimated using field replicate and laboratory replicate data.

4.11.1 Standard Operation Protocols (SOPs)

Each SOP should be a written detailed method accessible to each analyst. SOPs should be based on procedures developed by a standard-setting organisation such as Environment Canada, the U.S. EPA, the American Society for Testing and Material (ASTM), and American Public Health Association (APHA). Where methods are not well validated, it is recommended that the SOP be thoroughly referenced to the relevant literature and contain all the elements outlined in Canadian Association for Environmental Analytical Laboratories (CAEL) Inc. (1991). In-house validation of data should be appended to the SOP and should contain the quality assurance and quality control procedures. These include the types and frequencies of quality control samples to be analysed, expected levels of precision, accuracy, recovery, and method detection limits.

While chemical analysis procedures tend to be reasonably well documented, sampling procedures in general, and sampling design in particular, are often overlooked. Sampling error is usually a large component and often the dominant component of uncertainty in environmental measurements. Standard operating procedures that include field operations will help to reduce this uncertainty or at least ensure that it is quantified

4.11.2 Quality Control in the Field

This is the first stage of data collection. QA/QC procedures for the Fish Survey should be outlined during the development of the study design and should be followed precisely in order to maintain high quality data. While a QA/QC plan for field sampling can have many components, some of the main procedures are as follows:

- initiate and maintain communication with local government agencies (e.g., fishing licence, dates of fish collection, location of collection, etc.)
- all personnel involved in field sampling should have appropriate education and/or training
- all safety measures should be identified, understood and adhered to
- fish collection methods and equipment should be appropriate for the specific water-body and fish species.
- habitat descriptions including possible modifying factors (water depth and current, dissolved oxygen, temperature, substrate classification, evidence of pollution [discoloration, odour, residues], salinity, conductivity, etc.)
- date and time of collection
- collection methods need to be consistent throughout study
- location of sampling stations and fish collection areas documented; photograph collection location
- record of the number of fish species and incidental species caught per collection stations
- estimate of catch per unit effort
- samples from fish (e.g., ovaries, age structures,) should be placed in appropriate containers
- suitable preservatives/fixatives (e.g., ovaries - frozen or formalin,) should be used
- all samples should have appropriate labelling
- all measurements should be taken using appropriate equipment of acceptable accuracy and precision (this should be documented)
- instruments should be calibrated and maintained in good working order (records and methods should be available).

Table 4-7: Example of quality assurance checklist for aquatic biology studies (check or enter footnote number to provide details or to describe deficiencies.)

Quality Management

______ Standard operating procedures (SOP) available (sampling, processing and QA methods, revisions).
______ Quality management plan for institution (may be part of SOP).
______ Responsibilities, authorities, qualifications defined for each position.
______ Quality assurance officer with adequate authority for corrective action.
______ Data quality objectives for project (sensitivity, accuracy, precision, total uncertainty).

Quality Assurance Functions Performed

______ Review of QC sample results (permanent record, replicates, exchanges, verifications).
______ Performance assessment review (blind samples).
______ Knowledge testing and training sessions (field and laboratory).
______ Review of test sample data (transcription and logic checks performed).
______ Report approval mechanism (signatures).
______ Quality assurance reports (corrective actions indicated).
______ Inter-laboratory and accreditation studies (record of participation).

Field Operations

______ Preliminary sampling (return for effort, environmental strata, variability).
______ Sampling design (based on objectives, adequate replication).
______ Instrument calibration and maintenance (records, methods available)
______ Crew training and evaluation (video records and comparisons).
______ Sampling equipment (appropriate, consistent, cleaning and maintenance records).
______ Collection, preservation, shipping, storage (methods, adequate labels, custody records).
______ Additional effort checks (confirm little gain of information with extra effort).

Sample Processing, Quality Control, and Reporting

______ Sorting and subsampling (methods, adequate records, error estimates).
______ Sorting verification (confirm few organisms overlooked).
______ Taxonomy and enumeration (appropriate keys, records, reference collection).
______ Replicate processing (within and between personnel, error estimates).
______ Exchange samples (between laboratories, comparability assessed).
______ Taxonomic verification (by recognised experts).
______ Fish age verification (by recognised experts).
______ Archives (for samples, voucher specimens).
______ Data handling and reporting (entry checks, missing value codes, methods, QC data).

Details/Deficiencies

Footnote #	Description

- detailed field notes should be maintained in a bound notebook
- chain-of-custody forms and appropriate shipping and storage procedures should be used

4.11.3 Quality Control in the Lab

Although much of the survey information is collected while in the field, variables such as fecundity, egg size and age are usually determined later in the laboratory. With each measurement, the primary concern of the laboratory QA/QC program is to ensure consistency (precision) and accuracy of the data. The following issues should be considered as part of the measurement procedures:

- all personnel involved in sample processing and analyses should have appropriate experience, education and/or training
- measurements should be conducted using recognised protocols and methods (these should be documented) all instruments used should be properly calibrated and maintained (records, methods available)
- record of fish measurements recorded for each fish (target species)
- record of external lesions, tumours and parasites, etc.
- record of evidence of parasitism
- fecundity data including methods and subsampling precision (if applicable)
- ageing data including methods and independent confirmation of estimates
- records should be kept describing the sample, measurement, and responsible personnel. If possible, a minimum number of individuals should conduct a particular measurement to maintain consistency and reduce measurement error (especially for age determination)
- if subsampling is necessary (e.g., fecundity, egg size), information describing the efficiency and accuracy of the subsampling technique should be documented. This information should also be used to calculate appropriate correction or scaling factors (if needed) to minimise possible differences in methods and efficiency.
- all data should be verified. For example, measurements such as fecundity and egg size should be replicated to ensure precision and accuracy. A recognised expert should verify estimates of age.
- literature and taxonomic keys used for fish identification should be documented
- maintain detailed sample processing and laboratory notes in a bound notebook.

4.11.4 Data Analysis

QA/QC concerns regarding data analysis include data verification and validity, repeatability and robustness of statistical analyses, and rigour and defensibility of analyses. For the most part, validation and verification of data depends on the success of QA/QC procedures during field sampling, sample processing, and laboratory analyses (see above). However, there are other considerations regarding the data verification and analyses:

- conduct screening techniques to identify possible transcription errors, outliers and other "suspicious" data points,
- tabular summaries of the general descriptive statistics (sample size, mean, median, minimum, maximum and standard deviation) of fish measurements,
- results of assessing assumptions of normality and homogeneity of variance,
- record of transformation used,
- parameter estimates of variability (ANOVA MSE, ANCOVA MSE, SD for age-to-maturity),
- calculated sample size requirements for each parameter,

- summary of adherence to data quality objectives, standard operating procedures and identification of any QA/QC problems incorporate considerations related to laboratory and field QA/QC,
- to allow reproduction of analyses and results, provide all raw data in an Appendix and archive computer data files for an approved period of time after the analyses are published in a report,
- document in detail a clear description of the methods used for analyses,
- verify that statistical software packages used produce the same output and results as other packages,
- evaluate the robustness of the analyses, i.e. the results and conclusions should be similar,
- regardless of whether outliers are included or excluded, transformations are used or not, etc. The objective is to ensure that results are not a function of some manipulation or assumption prior or during analyses,
- maintain detailed notes regarding the analyses of the survey data.

4.11.5 Data treatment

Upon completion of the field and laboratory measurements (including the appropriate QA/QC procedures), a data matrix with the location identifier (area), variables in columns, and observations in rows operates as the fundamental working unit. An electronic copy of the data is submitted to the AO. The initial step in data analysis is a check for transcription errors. Failure to identify these errors can invalidate further analyses. Assuming the data have been entered correctly, data should be summarised, screened for erroneous values and outliers, assessed for normality and transformed.

Visual screening techniques such as Box and Whisker Plots, Normal Probability Plots, and Stem-and-Leaf Diagrams can be used to identify extreme values (true outliers and/or data entry errors). Most statistical software packages provide data summary modules capable of generating appropriate summary statistics and graphics. It is also beneficial to summarise the data by calculating the mean and standard deviation (SD) of each variable separately for the reference and exposure areas. These summary statistics are usually needed for presentation and aberrantly high or low values can indicate errors. Extreme values or outliers should not be removed from the data set (unless they are obvious sampling, measurement or data entry errors) (Grubbs 1969; Green 1979) because this will result in the loss of statistical power to the fish survey . Instead, extreme values should be identified in the report and the influence of the extreme value(s) on the results should be determined by re-analysing the data without the extreme value.

4.12 References

Andrews, S. and W. R. Parker. 1999. An environmental quality evaluation of Pictou Harbour, Nova Scotia, using caged bivalves, *Mytilus edulis*. Environment Canada, Environmental Protection Branch, Atlantic Region Surveillance Report. 30 p + appendices. (In progress).

Aquatic Effects Technology Evaluation (AETE). 1999. Report of Selected Technologies for Cost-effective Environmental Monitoring of Mine Effluent Impact in Canada. AETE synthesis Report. Prepared by ESG International Inc. for Aquatic Effects Technology Evaluation (AETE) Program, CANMET, Natural Resources Canada. Draft, February 1999.

Aquatic Effects Technology Evaluation (AETE). 1998. Technical Evaluation of Fish Methods in Environmental Monitoring for the Mining Industry in Canada. Prepared by EVS Environment consultants for Aquatic Effects Technology Evaluation (AETE) Program, CANMET, Natural Resources Canada. Draft, July 1998.

Aquatic Effects Technology Evaluation (AETE). 1997. Technical Evaluation of Molluscs as a biomonitoring tool for the Canadian Mining Industry. Prepared by Robin Stewart and Diane F. Malley for Aquatic Effects Technology Evaluation (AETE) Program, CANMET, Natural Resources Canada. March 1997.

Alldredge, J.R. 1987. Sample size for monitoring of toxic chemical sites. Envir. Monitoring Assess. 9:143-154

AQUAMIN Working Group 1. 1994. Criteria to Evaluate Information for AQUAMIN. Assessment of the Aquatic Effects of Mining in Canada: AQUAMIN, prepared for the AQUAMIN steering Group. Working Document, 8pp.

ASTM. 1988. Draft Standard Practice for Generation of Environmental Data. Committee E-47 on Biological Effects and Environmental Fate. Subcommittee E47.05 on Quality Assurance. American Society for Testing and Materials, Philadelphia.

ASTM. 1986. Manual on Presentation of Data and Control Chart Analysis. Committee E-11 on Statistical Methods. ASTM Special Technical Publication 15D. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1985. Standards on Precision and Bias for Various Applications. 2nd ed. American Society for Testing and Materials, Philadelphia, PA.

Bartlett, J.R., P.F. Randerson, R. Williams and D.M. Ellis. 1984. The use of analysis of covariance in the back-calculation of growth in fish. J. Fish. Biol. 24: 201-213

Bernstein, B.B. and J. Zalinski. 1983. An optimum sampling design and power test for environmental biologists. J. Envir. Mgmt. 16:35-43.

Bloom, N. S. 1992. On the chemical form of mercury in edible fish and marine invertebrate tissue. Can.J.Fish.Aquat.Sci. 49:1010-1017.

Bolger, T. and P.L. Connolly. 1989. The selection of suitable indices for the measurement and analysis of fish condition. J. Fish. Biol. 34:171-182

Burt, A., Kramer, D.L., K. Nakasuru and C. Spry. 1988. The tempo of reproduction in *Hyphessobrycon pulchripinnis* (Characidae), with a discussion on the biology of "multiple spawning". Environ. Biol. Fish. 22: 15-27.

CAEAL (Canadian Association of Environmental Analytical Laboratories). 1991. Code of Practice and QA Manual for Laboratory Analysis of Sewage Treatment Effluent in Support of the MISA Program; Draft report prepared for CAEAL and the Ontario Ministry of the Environment by Zenon Environmental Laboratories.

Clarkson, T.W. 1994. The toxicology of mercury and its compounds. Chapter VIII.1. In: Mercury pollution: Integration and synthesis, C.J. Watras and J.W. Huckabee, eds. Lewis Publishers, USA. pp. 631-641.

Coad, B.W. 1995. Encyclopædia of Canadian Fishes. Canada Museum of Nature and Canadian Sportfishing Productions Inc. 928 pp.

Cone, R.S. 1989. The need to reconsider the use of condition indices in fisheries science. Trans. Amer. Fish Soc. 118:510-514.

COSEWIC (Committee on the Status of Endangered Wildlife in Canada). 1998. Species presently registered as rare, threatened or endangered in Canada. http://www.cosewic.gc.ca/, COSEWIC, Canadian Wildlife Services.

Courtenay, S.C., W.R. Parker and G.P. Rawn. 1998. Proceedings of a Workshop to Assess Alternatives to the Fish Survey Component of the Environmental Effects Monitoring Program for Canadian Pulp and Paper Mills. Canadian Technical Report of Fisheries and Aquatic Sciences 2233: 8 + 100p.

Courtenay, S.C., K.R. Munkittrick, H.M.C Dupuis, R. Parker and J. Boyd. 2002. Quantifying impacts of pulp mill effluent on fish in Canadian marine and estuarine environments: problems and progress. Water Qual. Res. J. Canada 37(1) 79-99.

Culp, J.M. and C.L. Podemski. 1995. Design and application of a novel stream microcosm system for assessing effluent impacts to large rivers. *In* M.R. Servos, K.R. Munkittrick, J.H. Carey, and G.J. Van Der Kraak (eds.), Fate and Effects of Pulp and Paper Mill Effluents. St. Lucie Press, Boca Raton, Florida.

Culp, J.M., C.L. Podemski, K.J. Cash, and R.B. Lowell. 1996. Utility of field based artificial streams for assessing effluent effects on riverine ecosystems J. of Aquatic Ecosystem Health 5: 117-124.

Draper, N.R. and H. Smith. 1981. Applied regression analysis. 2nd ed. John Wiley & Sons, New York.

Dubé, M.G., D.L. MacLatchy, and M-K.A. Elliot. 1997. Effects of different effluent streams from a bleached kraft pulp mill on Fundulus heteroclitus: 2) After reverse osmosis treatment of a clean condensate. University of New Brunswick (Saint John), unpublished data.

Dubé, M.G., J.M. Culp and J.A. Smith. 1998. Development and testing of a mesocosm technology for environmental effects monitoring (EEM) methods application. *In* Courtenay, S.C., W.R. Parker and G.P. Rawn. 1998. Proceedings of a Workshop to Assess Alternatives to the Fish Survey Component of the Environmental Effects Monitoring Program for Canadian Pulp and Paper Mills. Canadian Technical Report of Fisheries and Aquatic Sciences 2233: 8 + 100p.

Dux, J.P. 1986. Handbook of Quality Assurance for Analytical Chemistry Laboratory. Van Nostrand Reinhold Co., New York.

Environment Canada and Department of Fisheries and Oceans (EC/DFO) 1995. Further guidance for the adult fish survey for aquatic environmental effects monitoring related to federal *Fisheries Act* requirements. EEM 1. Environment Canada, Ottawa, Ont. 72 pp.

Elder, J.F. and J.J. Collins. 1991. Freshwater molluscs as indicators of bioavailability and toxicity of metals in surface-water systems. Rev. Environ. Contam. Toxicol. 122: 37-79.

Environment Canada. 1998. Pulp and Paper Technical Guidance for Aquatic Environmental Effects Monitoring. EEM/1998/1.

Environment Canada. 1997. Fish Survey Expert Working Group Report. EEM/1997/6.

Environment Canada. 1991. Quality Assurance Guidelines for Biology in Aquatic Environmental Protection. Prepared by Dr. D.Hart, Beak Consultants Limited for Dr. A.S.Y. Chau, National Water Research Institute, Burlington, Ontario.

FAU Update Issue 94-1, Ontario Ministry of Natural Resources Fisheries Assessment Unit Newsletter. 1994. Nearshore community index netting (NSCIN): indexing the abundance of the warmwater fish community. Lake Simcoe Fisheries Assessment Unit, Sutton West, Ont. 2pp.

FAU Update Issue 94-2, Ontario Ministry of Natural Resources Fisheries Assessment Unit Newsletter. 1994. Spring littoral index netting (SLIN): indexing the abundance of the coldwater fish community. Muskoka Lake Fisheries Assessment Unit, Bracebridge, Ont. 2pp.

Foran, J. and S. Ferenc. 1999. Multiple Stressors in Ecological Assessments. SETAC Press, Pensacola, Florida. In press.

Fox, G.A. 1991. Practical causal inference for ecoepidemiologists. J. Toxicol. Environ. Health 33: 359-373.

Fritz, E.S., W.H. Meredith and V.A. Lotrich. 1975. Fall and winter movements and activity level of the mummichog *Fundulus heteroclitus*, in a tidal creek. Chesapeake Sci. 16:211-215.

Gibbons, W.N., K.R. Munkittrick and W.D. Taylor. 1998a. Suitability of small fish species for monitoring the effects of pulp mill effluent on fish. 1. Response of spoonhead sculpin exposed to bleached kraft mill effluent. Environ. Toxicol. Chem. In Press, NWRI Publication number 97-26a

Gibbons, W.N., K.R. Munkittrick and W.D. Taylor. 1998b. Suitability of small fish species for monitoring the effects of pulp mill effluent on fish. 2. Comparison of trout perch response to white sucker exposed to bleached kraft mill effluent. Environ. Toxicol. Chem. In Press, NWRI Publication number 97-26b

Gibbons, W.N., M.D. Munn and M.d. paine. 1993. Guidelines for monitoring benthos in freshwater environments. Prepared for Environment Canada, North Vancouver, B.C. by EVS Consultants, North Vancouver, B.C. 81pp.

Gibbons, W.N., M.D. Paine, W.M. Gibson, G.A. Vigers and K. Kingston. 1992. 1991 operational monitoring survey of the Lesser Slave River. Volume III. Prepared for Slave Lake Pulp Corporation, Edmonton, Alberta by EVS Consultants, North Vancouver, B.C. 63pp.+Tables + Figures + App.

Goldstein, R. 1989. Power and sample size via MS/PC-DOS computers. Amer. Statist. 43:253-260.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley & Sons, Toronto. 257 pp.

Gray, M.A., A.R. Curry, K.R. Munkittrick. 2002. Non-lethal sampling methods for assessing environmental impacts using a small-bodied sentinel fish species. Water Qual. Res. J. Canada 37(1) 195-211.

Green, R.H. 1989. Power analysis and practical strategies for environmental monitoring. Envir. Research 50:195-205.

Green, R.H. 1993. Application of repeated measures designs in environmental impact and monitoring studies. Australian Journal of Ecology 18:81-98.

Grubbs, F.E. 1969. Procedures for detecting outlying observations in samples. Technometrics 11:1-22.

Halbach, S. 1985. The octanol/water distribution of mercury compounds. Arch. Toxicol. 57: 139-141.

Heins, D.C. and F.G. Rabito Jr. 1986. Spawning performance in North American minnows: direct evidence of the occurrence of multiple clutches in the genus *Notropis*. J. Fish. Biol. 28: 343-357.

Hodson, P.V., B.R. Blunt, and D.M. Whittle.1984. Monitoring lead exposure of fish. Pg. 87-98. *In:* Contaminant Effects of Fisheries. Cairns, V.W., P.V. Hodson, and J.O. Nriagu. John Wiley and Sons. New York, New York.

Hunn. J.B. 1988. Field assessment of the effect of contaminants on fishes. U.S. Department of the Interior, Fish and Wildlife Service Biological Report 88(19). 27 pp.

International Standards Organization (ISO). 1992. Sensory Analysis: Vocabulary. International Standards Organization, Geneva, Switzerland. ISO 5492.

Jackson, D.A., H.H. Harvey and K.M. Somers. 1990. Ratios in aquatic sciences: statistical shortcomings with mean depth and morphoedaphic index. Can. J. Fish. Aquat. Sci. 47:1788-1795.

Jenkins, R.E. and N.M. Burkhead. 1993. Freshwater fishes of Virginia. American Fisheries Society, Bethesda, Maryland. 1079 pp.

Leblanc, J, C.M. Coulliard and J.F. BrÊthes. 1998. Modifications of the reproductive period in mummichog (Fundulus heteroclitus) living downstream from a bleached kraft pulp mill in the Miramichi Estuary, New Brunswick, Canada. Can. J. Fish. Aquat. Sci. 54:2564-2573.

Leblanc, J. et C. M. Couillard. 1995. Description de la période de reproduction d'un poisson sentinelle: le choquemort (Fundulus heteroclitus) de l'estuaire de la Miramichi Rapp. tech. can. sci. halieut. aquat. 2057: vii + 32p.

Kilgour, B.W., K.M. Somers and D.E. Matthews. 1998. Using the normal range as an ecological criterion for ecological significance in environmental monitoring and assessment. Ecoscience : in Press.

Lowell, R.B. 1997 Discussion paper on critical effect size guidelines for EEM using benthic invertebrate communities. Report to the Environmental Effects Monitoring Program. EEM/1977/9.

Major, M.A., D.H. Rosenblatt, and K.A. Bostian. 1991. The octanol/water partition coefficient of methylmercuric chloride and methylmercuric hydroxide in pure water and salt solutions. Environ. Toxicol. Chem. 10: 5-8.

MacMahon, R.F. 1991. Mollusca: Bivalvia, pp. 315-399. *In:* Ecology and classification of North American freshwater invertebrates. J.H. Thorp and A.P. Covich, (eds.). Academic Press, New York.

McMaster, M.E., G.J. Van Der Kraak and K.R. Munkittrick. 1996. An epidemiological evaluation of the biolochemical basis for steroid hormone depressions in fish exposed to industrial wastes. J. Great Lakes Res. 22: 152-171.

Minns, C.K. 1995. Allometry of home range size in lake and river fishes. Can. J. Fish. Aquat. Sci. 52: 1499-1508.

Munkittrick, K.R. and M.E. McMaster. 1999. Assessment of multiple stressors in aquatic ecosystems by directed assessment of cumulative effects using fish populations. Submitted manuscript.

Munkittrick, K.R. 1992. A review and evaluation of study design considerations for site-specificity in assessing the health of fish populations. J. Aquat. Eco. Health 1:283-292.

Munkittrick, K.R. and D.G. Dixon. 1989. Use of white sucker (*Catostomus commersoni*) populations to assess the health of aquatic ecosystems exposed to low-level contaminant stress. Can. J. Fish. Aquat. Sci. 46: 1455-1462.

Muramota, S. 1981a. Vertebral malformation in fish induced by decalcification from the bones with cadmium. Verh. Internat. Verein. Limnol. 21:1309-1313.

Muramota, S. 1981b. Vertebral column damage and decrease of calcium concentration in fish exposed experimentally to cadmium. Environ. Poll. 24: 125-133.

Nelson, J.S. and M.J. Paetz. 1992. The fishes of Alberta. The University of Calgary Press, Calgary, Alberta.

NRCC. 1985. The role of biochemical indicators in the assessment of ecosystem health - their development and validation. Publ. No. NRCC 24371, Ottawa. 119 pp.

OECD (Organisation for Economic Cooperation and Development). 1989. Draft Council Decision-recommendation on Compliance with Principles of Good Laboratory Practice (restricted). Organisation for Economic Cooperation and Development, Paris.

OECD (Organisation for Economic Cooperation and Development). 1981. Decision of the Council Concerning Mutual Acceptance of Data in the Assessment of Chemicals. Annex 2. OECD Principles of Good Laboratory Practice. In: Guidelines for Testing of Chemicals. Organisation for Economic Cooperation and Development, Paris.

OMOE (Ontario Ministy of the Environment)). 1984. Principles of Control Charting. D.E. King and R.C. Ronan, Laboratory Services Branch, Data Quality Report Series. Ontario Ministry of the Environment, Rexdale, Ontario.

Paine, M.D. 1990. Life history tactics of darters (*Percidae: Etheostomatiini*) and their relationship with body size, reproductive behaviour, latitude and rarity. J. Fish Biol. 37: 473-488.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries research and management. Can. J. Fish. Aquat. Sci. 47:2-15.

Peters, R.H. 1983. The ecological implication of body size. Cambridge University Press, New York. 329 pp.

QAMS (quality assurance management staff). 1986. Development of Data Quality Objectives. Description of Stages I and II. Environmental Protection Agency, Washington, D.C. (Draft).

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Fish. Res. Bd. Can. Bull. 191.

Roberts, W.E. 1988. The sculpins of Alberta. Alberta Naturalist 18:121-127.
Saouter, E., L. Hare, P.G.C. Campbell, A. Boudou, and F. Ribeyre. 1993. Mercury accumulation in the burrowing mayfly *Hexagenia rigida* (Ephemeroptera) exposed to CH_3HgCl or $HgCl_2$ in water and sediment. Water Res. 27: 1041-1048.

Salazar, M.H. and Salazar, S.M. 1997a. Using caged bivalves for environmental effects monitoring at pulp and paper mills: Rationale and historical perspective. In: J.S. Goudey, S.M. Swanson, M.D. Treissman and A.J. Niimi. Proceedings of the 23rd Annual Aquatic Toxicity Workshop, October 1996, Calgary, Alberta. Canadian Technical Report of Fisheries and Aquatic Sciences No. 2144129-136.

Salazar, M.H. AND Salazar, S.M. 1997b. Using caged bivalves to characterize exposure and effects associated with pulp and paper mill effluents. Wat. Sci. Tech. 35(2-3):213-220.

Sandström, O. 1996. *In situ* assessments of the impact of pulp mill effluent on life-history variables in fish. pp. 449-457 In: Environmental fate and effects of pulp and paper mill effluents. M.R. Servos, K.R. Munkittrick, J.H. Carey and G. J. Van Der Kraak (eds.). St. Lucie Press, Delray Beach, Florida.
Schetagne et al. 1996.

Schetagne et Verdon. 1999.

Scott, W.B. 1967. Freshwater Fishes of Eastern Canada. University of Toronto Press, Toronto, Ontario, Canada. 137 pp.

Scott, W.B. and E.J. Crossman, 1973. Freshwater Fishes of Canada. Fish. Res. Board. Can. Bull. 184. Ottawa, Ontario, Canada. 966 pp.

Shuter, B.J. 1990. Population-level indicators of stress. Amer. Fish. Soc. Symposium 8:145-166.

Skalski, J.R. and D.H. McKenzie. 1982. A design for aquatic monitoring programs. J. Envir. Mgmt. 14:237-251.

Snedecor, G.W. and W. G. Cochran. 1971. Statistical Methods. Sixth Edition. Iowa State University Press.

Snedecor, G.W. and W.G. Cochran. 1980. Statistical Methods, 7th ed. Iowa State University Press, Ames, Iowa. 507 p.

Swanson, S.M. 1993. Wapiti/Smoky river ecosystem study. Prepared for Procter and Gamble Ltd./Weyerhaeuser Canada Ltd., Grand Prairie, AB by Sentar Consultants Ltd., Calgary, AB. 176 pp.

Taylor, J.K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers Inc., Chelsea, Michigan.

Underwood, A.J. 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecological Applications 4(1):3-15.

Underwood, A.J. 1997. Experiments in ecology: their logical design and interpretation using analysis of variance. Cambridge University Press. United Kingdom. 504 pp.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell and C.E. Cushing. 1980. The river continuum concept. Can. J. Fish. Aquat. Sci. 37:130-137.

Vetemaa, M., L. Forlin and O. Sandstrom. 1997. Chemical industry effluent impacts on reproduciton and biochemistry in a North Sea population of viviparous blenny (*Zoarces viviparous*). J. Aquat. ecosys. Stress & Recovery 6:33-41.

Zar, J.H. 1984. Biostatistical Analysis, 2nd ed. Prentice-Hall, Englewood Cliffs, NJ.

Table of Contents

5. Benthic Invertebrate Community Survey *5-1*

5.1 Program Overview and Objectives of the EEM Phases *5-1*
5.1.1 Site Characterization 5-1
5.1.2 First EEM Study 5-2
5.1.3 Subsequent EEM Studies 5-3
5.1.3.1 Periodic Monitoring 5-3
5.1.3.2 Focussed Monitoring 5-3
5.1.3.3 Investigation of Cause 5-4
5.1.4 Benthic Invertebrate Community Survey in Marine and Estuarine environments 5-5

5.2 Study Design Considerations for Field Monitoring Programs *5-6*
5.2.1 Review and Summarizing Historical Information 5-6

5.3 Consideration of Site-Specific Conditions for Study Design Exercises *5-8*

5.4 Monitoring Frequencies *5-12*

5.5 Standard Nomenclature and Reporting Formats for Field Locations *5-13*
5.5.1 Standard Nomenclature 5-13
5.5.2 Reporting of Field Station Positions 5-18

5.6 Sampling Program Designs *5-18*

5.7 Alternative Monitoring Methods *5-24*

5.8 Selection of Reference and Exposure Areas *5-37*
5.8.1 General guidance for Reference Area Selection 5-37
5.8.2 Habitat-Specific Allocation of Reference and Exposure Areas 5-38
5.8.3 Reference and Exposure Area Consideration for the EEM program phases5-40

5.9 Selection of Ecologically Relevant Habitats *5-41*
5.9.1 General Guidance for Habitat Selection 5-41
5.9.2 Habitat Considerations for the EEM program phases 5-42

5.10 Selection of Ecologically Relevant Sampling Seasons *5-46*
5.10.1 General Guidance for Sampling Season Selection 5-46
5.10.2 Sampling Season Considerations for the EEM program phases 5-47

5.11 Statistical Considerations for Study Design *5-47*
5.11.1 Setting Effect Sizes for determination of sampling effort 5-48
5.11.2 Selecting α and β 5-50
5.11.3 Determination of Sampling Effort using Power Analysis for ANOVA (and Regression/ANCOVA) designs 5-52
5.11.4 Determination of Sampling Effort for RCA designs 5-54
5.11.5 Determination of Sampling Effort for field sub-sampling 5-55
5.11.6 The Use of ordination probability ellipses for RCA designs 5-58

5.12 Benthic Invertebrate Taxonomy ***5-61***
5.12.1 Taxonomic Level of Identification 5-61
5.12.2 Reference Collections 5-62

5.13 Supporting Environmental Variables ***5-63***

5.14 Integrating Benthic Invertebrate and Sediment Programs ***5-65***
5.14.1 Review of the Supporting Environmental Variables for Subsequent Monitoring 5-66

5.15 Other Supporting Environmental Variables ***5-68***

5.16 Data Assessment and Interpretation ***5-68***
5.16.1 Data Handling Methods 5-68
5.16.1.1 QA/QC for Data Input and Verification 5-68
5.16.1.2 Dealing with outliers 5-69
5.16.2 Unknown and Immature Organisms 5-69
5.16.3 Data Reductions and Transformations 5-70

5.17 Data Reporting Guidelines ***5-71***

5.18 Summarizing Monitoring Results ***5-72***
5.18.1 Calculation and Reporting of Benthic Invertebrate Community Descriptors5-73

5.19 Evaluation of Results ***5-78***
5.19.1 Is there an effect on the benthic invertebrate community? 5-78
5.19.2 What is the next MM EMM Step? 5-79
5.19.3 Summary of biological monitoring frequencies for the benthic invertebrate community after data has been assessed 5-84

5.20 EEM program options after an effect has been established ***5-85***
5.20.1 Is the effect mine related? 5-85
5.20.2 Is the magnitude and geographical extent known? 5-86
5.20.3 Is the cause of the effect known? 5-86

5.21 Methods for Benthic Invertebrate Monitoring ***5-87***
5.21.1 Field Methods 5-87
5.21.1.1 Sampler Mesh Sizes 5-87
5.21.1.2 Sampling Equipment 5-89
5.21.1.3 Specimen Preservation 5-96
5.21.2 Methods for Supporting Environmental Variables 5-96
5.21.3 QA/QC for Benthic Invertebrate Field Operations 5-97

5.22 Laboratory Methods ***5-98***
5.22.1 Sample Sorting 5-98
5.22.2 Laboratory Sub-sampling 5-100
5.22.3 Taxonomy and Reference Collections 5-103
5.22.4 QA/QC for Benthic Invertebrate Laboratory Operations 5-108

5.23 Tools for Alternative Monitoring Methods and/or Investigation of Cause5-109
5.23.1 Use of Weight-of-Evidence Approaches to Establish Cause of Effect(s) 5-109
5.23.2 On-Site Community Bioassays Using Mesocosms as an Alternative

Monitoring Method or Bioassessment Tool 5-112
5.23.3 Methods for Determining Biomarkers and Metal Body Burdens 5-114
5.23.4 Lethal and Sublethal Toxicity Tests 5-116
5.23.5 Analysis of Sediment Cores for Historical Trends 5-119

5.24 Other Benthic Invertebrate Measures and Organisms........ 5-119

5.25 References (list incomplete)........ 5-131

List of Tables

Table 5-1: Confounding Factors (see Hauer and Lamberti (1996) for further discussion). 5-10
Table 5-2: Alternative Monitoring Methods for benthic invertebrate community surveys 5-11
Table 5-3: Recommended benthic invertebrate community sampling program designs.5-25
Table 5-4: Potential outcomes from statistical analyses. (✓= no error, conclusions made are correct) 5-51
Table 5-5 Sample sizes to detect difference of 2 SD 5-54
Table 5-6: Recommended supporting explanatory variables for freshwater benthic invertebrate habitats. 5-65
Table 5-7: Recommended supporting explanatory variables for marine/estuarine benthic invertebrate habitats. 5-67
Table 5-8a: Summary criteria for recommended samplers in freshwater 5-92
Table 5-8B: Summary criteria for recommended samplers in marine/estuarine habitat.... 5-95
Table 5-9: Taxonomic Keys for benthic invertebrate taxonomic identification in freshwater environments. 5-104
Table 5-10: Recommended level of taxonomic precision for benthic invertebrates in marine environment 5-106
Table 5-11: List of marine and estuarine taxonomic benthic invertebrate keys for Canada5-107
Table 5-12: Formalized set of causal criteria forming part of a weight-of-evidence approach for assessment of mining effluent effects (Culp et al., 2000, Lowell et al., 2000). 5-111

List of Figures

Figure 5-1: Decision Tree for the Assessment of Existing or Historical Information.... 5-9
Figure 5-2 Examples of AREA, REPLICATE STATION and FIELD SUB-SAMPLE spatial scales for a basic Control/ Impact design. 5-17
Figure 5-3a: Control/Impact designs for simple freshwater rivers and streams or for homogeneous estuarine habitat. 5-26
Figure 5-3b: Modified C/I Design with downstream reference area for stream, river or estuaries 5-27

Figure 5-3c: Focussed Monitoring design example (currently under review) 5-28
Figure 5-4a&b: Multiple control/impact designs for freshwater rivers and streams with two reference areas. 5-29
Figure 5-4c: Freshwater, geographically homogeneous lakes and marine bays or inlets (currently under review) 5-30
Figure 5-5A: Simple gradient designs for freshwater rivers or streams, and estuaries 5-31
Figure 5-5b: Simple gradient designs for lake or coastal mines situated in narrow bays or fjords. 5-32
Figure 5-6: Radial gradient design for lake or coastal situations 5-33
Figure 5-7a&b: Multiple Gradient Designs (currently under review)........................... 5-34
Figure 5-8: Reference Condition approach.. 5-35
Figure 5-9: Impairment stress levels derived for reference sites in HMDS ordination space. .... **Error! Bookmark not defined.**
Figure 5-10: Decision Trees for Data Assessment 5-81
Figure 5-11: Benthic invertebrate sorting and sub-sampling protocol, which would be applicable for samples with detrital material and large numbers of small organisms. 5-102

5. Benthic Invertebrate Community Survey

The objective of the Metal Mining EEM (EEM) program is to evaluate the effects of mine effluent on the aquatic environment, specifically fish habitat, fish, and the use by man of fish. To evaluate the effect of mine effluent on fish habitat the following question is answered:

Has the mine effluent had an effect on the benthic invertebrate community?

This chapter provides technical guidance recommended for the design and implementation of the benthic invertebrate community survey of the program. This chapters objective is to provide sufficient, detailed guidance for the program and study design, and to direct knowledgeable study planners to the appropriate literature for further information.

5.1 Program Overview and Objectives of the EEM Phases

This section summarizes the objective, frequency and study design elements for each phase of the program in relation to the benthic invertebrate community survey. The overall framework of the EEM program is presented in Chapter 1 and Figure 1.1.

5.1.1 Site Characterization

The objective of Site Characterization is to obtain and summarize sufficient physical, chemical and biological data to be able to design an effective EEM program. For more information regarding Site Characterization, refer to Chapter 3.

Site Characterization information useful for the benthic invertebrate community survey includes:

1) Identification of potential reference and exposure areas;
2) General habitat mapping and classification within these areas; and,
3) Summarization of historical or existing benthic invertebrate data.

5.1.2 First EEM Study

The objective of the first EEM study is to gather and summarize sufficient benthic invertebrate community data to be able to answer the following question:

> ***Has the mine effluent had an effect on the benthic invertebrate community within the aquatic receiving environment most proximate to the effluent discharge?***

Answering this question can be done either using historical monitoring data or by conducting a benthic invertebrate community survey (Initial Monitoring). If a mine chooses to use historical monitoring data, the mine will submit a report to the authorization officer within one year of the coming into force of the EEM sections of the MMER. This report would also outline any historical data the mine had on the fish or fish tissue. Subsequently, the first EEM study design for a mine using historical data would be submitted within two years of the coming into force of the EEM sections of the MMER.

If the mine has no historical data, or the data are insufficient to determine if the mine effluent is having an effect on the benthic invertebrate community, fish population or fish tissue, the first EEM study design is submitted within one year of the coming into force of the EEM sections of the MMER.

The study design for the first EEM study should include:

1. One ecologically relevant season;
2. Sampling in reference and high exposure areas (e.g., area most proximate to effluent discharge);
3. Sampling in ecologically relevant habitat(s);
4. One of five site-specific sampling designs (Table 5.3);
5. Site-specific supporting variables; and,
6. Standardization of field and laboratory methods.

The benthic invertebrate community descriptors which are calculated, reported, and used as endpoints in determination of effects that are reported in the interpretative report include:

Total invertebrate density
Taxon richness
Simpson's diversity index
Bray-Curtis index

Additional community descriptors which should be calculated and reported and can aid in data interpretation but are not required to be used as endpoints in the determination of effects include:

Evenness
Taxon (i.e., Family) density
Taxon (i.e., Family) proportion
Taxon (i.e., Family) presence/absence

5.1.3 Subsequent EEM Studies

5.1.3.1 Periodic Monitoring

The objectives of Periodic Monitoring are to confirm results, detect changes in conditions of the aquatic receiving environment, and to allow for the generation of longer term trend data (i.e., temporal patterns).

Determination of the appropriate survey frequency is dependent on results from previous monitoring. A decision tree is provided in section 5.19 to aid in decisions regarding frequency.

The study design for Periodic Monitoring should include:

1) The same study and sampling design elements as the first EEM study (Initial Monitoring or historical monitoring data);
2) A review of previous monitoring results to improve study designs (if necessary); and,
3) One of three sampling frequencies:
    - Confirmation Monitoring – survey in next appropriate season (but within next 24 months)
    - Surveillance Monitoring – survey within next 36 months
    - Minimal Monitoring – survey within 72 months.

5.1.3.2 Focussed Monitoring

The objective of Focussed Monitoring is to determine the magnitude and geographical extent of effects on the benthic invertebrate community that are related to mine effluent. Focussed Monitoring is conducted in the next appropriate sampling season. The interpretative report is submitted to the authorization officer within 24 months of the previous interpretative report being submitted.

Focussed Monitoring study designs should include:

1) Study and sampling design elements similar to previous monitoring, but with more extensive geographic coverage;
2) An evaluation of the adequacy of previously sampled reference areas due to the greater geographic extent of Focussed Monitoring. The new geographic extent of

Focussed Monitoring may include additional habitats and substrata, such as higher order streams or lakes, or marine/estuarine areas ranging from the intertidal to subtidal. If these new habitats were not represented in the reference area(s) that were used during previous monitoring, a re-assessment of the adequacy of these reference areas is recommended;

3) The sampling of additional ecologically relevant habitats, seasons or invertebrate life stages if this is appropriate for assessing the magnitude of the effect;
4) A consideration of other biotic indicators as tools to assess the magnitude of the effect if their use is appropriate and value-added. The list of optional indicators includes biomass and taxonomic composition of periphyton, phytoplankton, macrophyte or zooplankton communities; sampling of other invertebrate life stages, lower level invertebrate identification, invertebrate biomass, secondary production, additional sensitive habitats or seasons; and toxicity tests of sediment and water.

5.1.3.3 Investigation of Cause

The objective of Investigation of Cause is to determine the cause of a mine-related effect when previous monitoring has failed to provide an explanation of the cause.
Possible corrective actions may be considered but are outside the EEM program. Environment Canada, in consultation with Fisheries and Oceans Canada, has developed a "Guidance Document for Determining Follow-up Actions when Effects have been identified in Environmental Effects Monitoring (EEM)" (Environment Canada 2001). This document provides guidance to regulatory agencies, facilities and EEM practitioners for determining follow up actions when effects have been identified in EEM. It discusses the factors that need to be considered for determining follow up actions, and explains the roles and responsibilities of government, industry and stakeholders in this process.

Investigation of Cause monitoring is conducted in the next appropriate sampling season. The interpretative report is submitted to the authorization officer within 24 months of the previous interpretative report being submitted.

Investigation of Cause study designs should include:

- A description of what field and laboratory studies will be used to determine the cause of the effect.

A list of possible Investigation of Cause monitoring tools is provided in section 5.23. The tools selected should be readily available, cost-effective, recognized with primary literature documentation, and applicable to the EEM program.

Once the cause is established, Periodic Monitoring will resume every 36 months to ensure that effects are not changing.

5.1.4 Benthic Invertebrate Community Survey in Marine and Estuarine environments

The most important impacts from mining in marine and estuarine environments will result mainly from acid rock drainage (ARD), disposal of processing chemicals and surface or submarine tailings disposal, habitat alteration (Kline 1994) and smothering of sessile fauna and clogging from re-suspended particulate material (Ellis and Hoover 1990, Ellis et al. 1994). Although there are examples of high levels of leached metals which could potentially affect subtidal benthos (c.f. Black Angel mine, Greenland – Ellis et al. 1994), the buffering and sedimenting effects of seawater and sediment sulfides generated by sulfide reducing bacteria (Perry 1995) generally reduce the availability of soluble metals to deep marine benthos. Correlational relationships have been shown to occur between sediment particulate metals at mine sites and infaunal species richness (c.f. Rygg 1985, Burd and Ellis 1995). However, ARD can be a problem in intertidal areas (Kline 1994) around mines with surface resuspension of solids or waste rock dumping grounds, and in low-salinity estuarine areas and where tailings are discharged to river systems discharging to the sea. Early settled larval stages of benthic organisms are most vulnerable to metal toxicity (Kline 1994). Although Kline (1994) summarized the literature regarding research on toxic effects of tailings metals and mine processing chemicals on benthic marine organisms, he pointed out that the synergistic or interactive toxicity of such factors is still largely unknown. Kline (op. Cit.) also points out that benthic invertebrates have a low tolerance to metals compared to fish. Finally, the effects of habitat alteration on benthic communities in areas where mine tailings are deposited, has not been carefully studied and has been deemed unpredictable (c.f. Hesse and Reim 1993).

Most of the historical information on effects from metal mines discharging to the marine is from west coast temperate and subarctic mines in B.C. and Alaska which are already closed. A good literature synthesis of environmental effects and monitoring of these mines is given in Baer et al. (1992), Kline (1994), with case studies described by Poling and Ellis (1993) and Ellis et al. (1994). Those Canadian mines with submarine tailings discharge currently in operation in the arctic include Polaris and Nanasivik in the Eastern Arctic circle.

Most of the information in this chapter is relevant to both temperate and sub-arctic mines. However, there are a few special considerations which affect benthos and which should be taken into account for arctic mines. These are related to the unusual environmental conditions, which are strongly seasonal in nature. In particular, effects from freezing, ice scour, extended low light levels, and related low temperatures and unusual salinity gradients should be taken into account in designing benthic surveys. Experts also suggest that recolonization of benthic commercial species in arctic areas is probably considerably slower than in temperate areas (Ellis et al. 1994, Poling et al. 1993).

Experts on the EEM Benthic Invertebrate Sub-group working on this chapter gathered their expertise to provide guidance on benthic invertebrate community survey in freshwater. Specific recommendations for conducting monitoring in marine and estuarine environments were consecutively added and are inserted in TEXT BOXES to stress differences between marine and freshwater monitoring approaches.

It is recognized that more efforts will need to be focused to specifically address issues with regards to marine monitoring. The format adopted in this chapter will ease the updating of the technical guidance for these situations when more technical information will be available and assessed.

5.2 Study Design Considerations for Field Monitoring Programs

5.2.1 Review and Summarizing Historical Information

During Site Characterization (Chapter 3) any benthic invertebrate community data that were collected prior to the commencement of the EEM program, and that may be appropriate for mine site assessment, should be identified. The major consideration for assessing the appropriateness of previously collected information is whether it can determine if the mine effluent is having an effect on the benthic invertebrate community. Historical information may include long-term data sets for the mine, as well as other benthic invertebrate data from the ecoregion found in the primary literature, government reports, and reports from other mines or industries (Chapter 3). Historical data for the benthic component of EEM are acceptable without reference to time since collection providing these data identify the QA/QC measures that were taken to ensure validity of the data. Thus, the data should meet EEM objectives for site-specific assessment. Furthermore, the mine operating conditions (e.g., loadings of deleterious substances) and environmental conditions (e.g., locations of dams, bridges, discharge points) should not have changed significantly since the information was collected.

Any historical benthic invertebrate sampling data that have been collected from potential exposure, and/or reference areas should be summarized as follows:

Sampling program and methods information:

- geographical description of the area(s) and station(s)
- number of stations
- number of field sub-samples taken
- type and size of sampler(s) used
- mesh size(s) of sampler and sample reduction or laboratory sorting sieves
- date of sampling trip(s)

- habitat type(s)
- supporting environmental variables measured
- laboratory sorting procedures (e.g., was sub-sampling used, taxonomic identification level)
- QA/QC procedures

Benthic Invertebrate Community Descriptors

- Total invertebrate density
- Taxon (i.e., Family) richness
- Simpson's diversity index
- Bray Curtis index
- Evenness
- Taxon (i.e., Family) density
- Taxon (i.e., Family) proportion
- Taxa (i.e., Family) presence/absence

If possible, the benthic invertebrate data should be summarized and organized to fit into the EEM terminology for field locations as described in section 5.5 (i.e., area, station and field sub-sample). For details on the calculation of benthic community descriptors and for guidance on summarizing monitoring results see section 5.18. These descriptors can be used directly during the study design exercises (section 5.11) for the various EEM phases, or, if sufficient data are available, to proceed through the Decision Tree for the Assessment of Existing Information (Figure 5.1) or Data Assessment (Figure 5.11).

Assessing Data Quality and Quantity

In assessing the quality and quantity of historical data for the EEM program, the following questions should be addressed by reviewing Figure 5.11:

- Can the data be used to calculate the descriptive statistics and metrics listed above for appropriate reference and exposure areas?
- Is there sufficient data to perform statistical analyses on the data to answer the question, "Is there an effect?", for the required descriptive statistics and metrics?
- If the statistical analyses indicate that there is *no effect*, is the power ≥ 0.90?

Note that, if the conclusion from the historical data is that there is *no effect* on the benthic invertebrate community, it is recommended that the mine perform a statistical analyses and to demonstrate power before proceeding to the appropriate phase of Periodic Monitoring (see further guidance in section 5.19). However, if there is an effect from the effluent on the receiving environment, but there is insufficient data to perform statistical analyses to demonstrate this effect, the mine can propose to proceed directly to a later phase (e.g., confirmation monitoring - Periodic, Focussed Monitoring or Investigation of Cause) in the EEM program.

5.3 Consideration of Site-Specific Conditions for Study Design Exercises

The decision tree for the overall EEM program (Chapter 1, Figure 1.1) is used as a general reference that links the program phases of EEM. The decision tree is described in detail with the overall program objectives in Chapter 1. An additional decision tree is provided for the Benthic Program as an aid for reviewing and evaluating existing and Historical information (Figure 5.1). This decision tree poses a series of questions regarding site-specific conditions that may affect potential study design options. It also directs study planners to further information in tables or text sections providing general guidance for study designs and cases where confounding factors may be present (Tables 5.1 and 5.2).

This decision tree is intended for use at the onset of all study design exercises for either **Initial** or **Periodic** Monitoring phases of the EEM Benthic Invertebrate Component (i.e., it is applicable for new and subsequent surveys). Furthermore, this series of questions (or an appropriate subset of these) should also be applicable during study design exercises for **Focussed Monitoring** and **Investigation of Cause** as the presence of confounding factors can influence the design of those programs.

Figure 5-1: Decision Tree for the Assessment of Existing or Historical Information

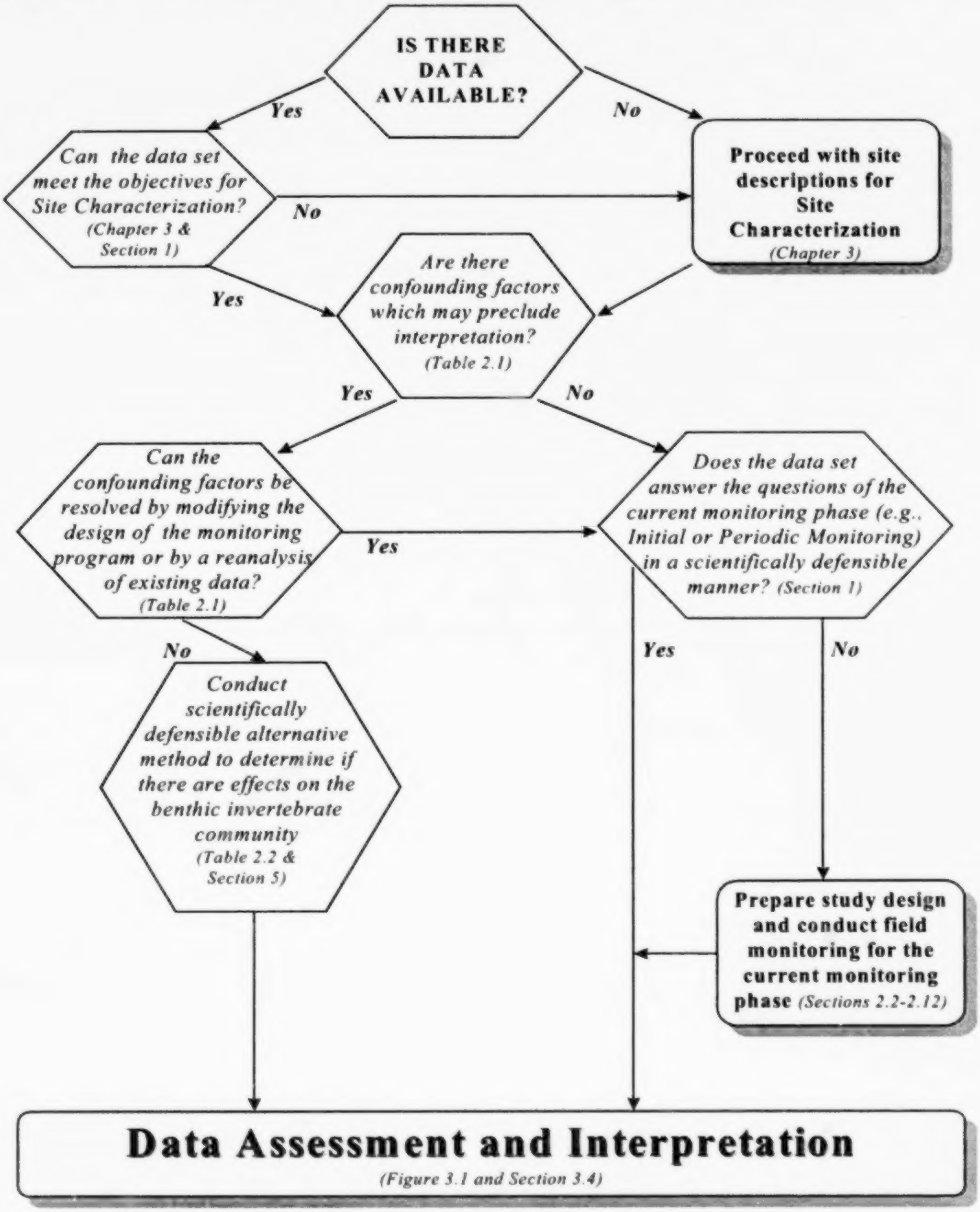

IS THERE DATA AVAILABLE?
Yes
No
Can the data set meet the objectives for Site Characterization? (Chapter 3 & Section 1)
No
Proceed with site descriptions for Site Characterization (Chapter 3)
Yes
Are there confounding factors which may preclude interpretation? (Table 2.1)
Yes
No
Can the confounding factors be resolved by modifying the design of the monitoring program or by a reanalysis of existing data? (Table 2.1)
Yes
Does the data set answer the questions of the current monitoring phase (e.g., Initial or Periodic Monitoring) in a scientifically defensible manner? (Section 1)
No
Yes
No
Conduct scientifically defensible alternative method to determine if there are effects on the benthic invertebrate community (Table 2.2 & Section 5)
Prepare study design and conduct field monitoring for the current monitoring phase (Sections 2.2-2.12)
Data Assessment and Interpretation (Figure 3.1 and Section 3.4)

Table 5-1: Confounding Factors (see Hauer and Lamberti (1996) for further discussion).

Note that, it is not the intention of the EEM program to mandate a single industry to investigate effects from multiple discharges on the benthic community.

Are there confounding factors which can be resolved by modifying the study design?

The interpretation of benthic community effects may be difficult if confounding factors exist within the study area. A careful review of historical or existing data and Site Characterization information of the study and sampling design elements can often resolve problems with confounding factors.
Four categories of such factors include:

Environmental variables: Environmental variables can confound the interpretation of benthic community effects if it is not possible to separate the effect of the mine effluent from the effects of changes in natural habitat variables. Augmenting the design to better characterize reference conditions with representation of all habitat types sampled may reduce the problem. This could include locating reference areas in adjacent or further afield watersheds, or by sampling additional reference areas (e.g., moving from a simple control-impact design to a more appropriate design, Figures 5.3-5.9 and Table 5.3). Some of the potentially confounding variables, that may be dealt with by applying more appropriate study and sampling designs, include depth gradients, substrate particle size, rapid effluent dilution, interannual and rare events, and seasonal and long term variability in physical characteristics such as temperature and flow regimes.

Multiple discharges or Historical effects: A potential for confounding effects exists if areas with varying levels of exposure to the mine effluent also have varying levels of exposure to other effluents or stressors, or habitat modifications from Historical factors such as dams or impoundments. If feasible, a change in sampling design by modifying sampling locations may reduce the problem. The collection of sediment cores may also be useful in depositional environments to resolve confounding factors resulting from Historical trends (see section 5.23.5 for further details).

Time of sampling: The time of year or the particular year of sampling may confound the interpretation of benthic community effects to effluent. This can be assessed by knowledge of the phenology of benthic community species (i.e., relation between climate and life history characteristics) and examination of data collected in previous years from reference areas.

Sampling methods: If standard methods (e.g., sampler types, mesh sizes, taxonomic levels) have not been used consistently within a study or in consecutive studies, this may obscure any benthic community response to the mine effluent. This confounding issue may occur when attempting to use Historical data from studies which were not designed to answer the EEM objectives for a particular mine site. It may be possible to examine the data in more detail and, if possible, convert the data to a similar comparable level (i.e., convert all taxonomic identification levels to a higher common level). However, in many cases a redesigned study ensuring that standard methods are consistently applied may be necessary to resolve these problems. Finally, if environmental or logistical conditions exist which preclude the safe and effective collection of samples, the applicability of alternative methods should be examined.

Table 5-2: Alternative Monitoring Methods for benthic invertebrate community surveys

Are there confounding factors which cannot be resolved?

If there are confounding factors that cannot be resolved through sampling design modifications, or a re-analysis of an existing data set, then the program proceeds to a consideration of alternative monitoring methods (see below).

Are there cost effective alternative monitoring methods?

The list below includes only a single alternative monitoring method and thus is not exhaustive. Mines may propose scientifically defensible additional approaches to determine effects on the benthic invertebrate community. Information regarding the applicability to the EEM program of the following alternative is briefly discussed in section 5.6 and 5.23. A more detailed guidance document on the use of this alternative method is also available (Chapter 12).

1) On-Site community bioassays using mesocosms as a bioassessment tool.

5.4 Monitoring Frequencies

The frequency of benthic invertebrate surveys for the EEM program is dependent on the results of previous monitoring. For all phases, an ***appropriate season*** is one that is both ecologically relevant *and* logistically feasible. The determination of ecological relevance should include consideration of the information in section 5.10. Logistic feasibility is based on several factors including:

1. field sampling feasibility,
2. safety of field personnel, and
3. the necessary timeframe for consultation such that sufficient time is allowed for redesign of the sampling program (if necessary).

First EEM Study

For **Initial Monitoring**, the **first** biological monitoring study, including the benthic survey, is to be completed, including the submission of the interpretative report (in an appropriate season) within 30 months of the mine being subject to Section 7 or the MMER. , The first study design is submitted within 12 months of the mine being subject to Section 7 of the MMER.

If a mine has **historical** fish, fish tissue or benthic invertebrate community data, the mine may submit a report to the authorization officer within 12 months of the mine being subject to Section 7 of the MMER. Subsequently, the first EEM study design would be submitted within 24 months of the mine being subject to Section 7 of the MMER. Therefore, for mines with historical monitoring data, the benthic survey would be completed, including submission of the first interpretative report, within 42 months of the mine being subject Section 7 of the MMER.

Subsequent EEM Studies

For **Periodic Monitoring** there are three sample frequency options that depend on results from previous surveys (Figure 5.11) and the objectives of the current Periodic Monitoring survey. Confirmation Monitoring requires a survey in the *next appropriate season,* meaning the Interpretative Report is submitted within **24 months** of previous Interpretative Report, *Schedule 5, S. 22(1)(a)*). Surveillance Monitoring requires a survey in an appropriate season within ***36 months***(including the submission of the Interpretative Report) of the previous interpretative *report (Schedule 5, S. 22(1)*). Minimal Monitoring requires a survey in an appropriate season within ***72 months*** of the submission of the previous Interpretative Report *(Schedule 5, S. 22(2)(b)*).

For **Focussed Monitoring** and **Investigation of Cause**, the survey or investigation is to be conducted within the ***next appropriate season,*** meaning the Interpretative Report must be submitted within 24 months of previous Interpretative Report *(Schedule 5, S. 22(2)(c)*).

5.5 Standard Nomenclature and Reporting Formats for Field Locations

5.5.1 Standard Nomenclature

Standardized definitions for sampling location nomenclature are essential to the EEM program because this will aid national or regional assessments of the EEM program. The following standardized terminology for sampling locations should be adopted and applied in a consistent and rigorous manner for all EEM studies with a benthic invertebrate community survey. Schematic representation of these terms is provided in Figure 5.2.

This section defines the terms, **Reference Area**, **Exposure Area**, **Replicate Station** and **Field Sub-Sample.** For the basic ANOVA study designs (i.e., Control/Impact, or Multiple Control/Impact, see section 5.6), where the objectives are to detect differences between or among areas, each reference or exposure area consists of a number of replicate stations (i.e., the *replicates* for an ANOVA). Each replicate station consists of a number of pooled field sub-samples. Similarly, gradient or reference condition study designs use the replicate station as the spatial scale of replication, with field sub-samples being collected as appropriate (see section 5.6 for a description of these approaches).

The concept of area as defined above, is not directly transferable to the gradient or RCA study designs. When using or developing a gradient or RCA study design, a balanced design with similar numbers of replicate stations located within reference and exposure areas is not the basis for comparisons. For example, in RCA study designs, exposure stations(s) are individually compared to a distribution of reference stations which represent appropriate reference conditions. For gradient designs, the lack of suitable reference or exposure "areas" may be the direct cause for selecting this study design and thus the ANOVA type terminology is not directly applicable for this approach. Detailed guidance on dealing with these study designs is included in Section 5.6, Sampling Program Designs.

Further guidance regarding the number of replicates and their allocation for different spatial scales and study designs is provided in section 5.11.3.

Field sub-sample

Field sub-samples will consist of individual area or time-limited collections of benthic invertebrates (e.g., one grab, core, cylinder, quadrat, kick or U-net sample). To ensure adequate spatial placement of field sub-samples within a station they should be collected in a random, or stratified-random pattern. For many of the statistical analyses used to assess effects in freshwater and marine environments (section 5.6), data from all field sub-samples within a station are pooled, providing a single value of each descriptor from each station.

Pooling of field sub-samples

The pooling of field sub-sample data can occur at several points in the monitoring program. At which point pooling occurs, depends on several factors including;

1) field sample processing and storage efficiency (e.g., are field storage jars large enough to contain pooled samples?),
2) laboratory sorting efficiency (e.g., is it more efficient to sort smaller samples?), and
3) the potential to address study design issues.

The first two factors result in an actual physical pooling of the samples, are considered logistical in nature and their applicability should be determined on a site-specific or method-specific basis. Note that once this physical pooling is done the potential information from individual sub-samples is lost. In regards to factor 3, if there is a need for additional information to address study design issues (e.g., to examine species area curves or field sub-sampling precision, see section 5.11.3 for more details), field sub-samples may be preserved and processed separately. The resultant un-pooled data are then available to address the study design issues and can subsequently be pooled electronically for the appropriate statistical analyses. Electronic pooling for the endpoints should be done in such a manner as to be equilvalent to results if field sub-samples were physically pooled. This is particularly important for the taxa richness endpoint. Example calculations for pooled station density and richness are included below.

For the density endpoint the value should be calculated as:

Density from pooled field sub-samples = (# in field subsample a + # in subsample b + # in subsample c)/total area of all field subsamples.

Note that, the resulting number is also the same as calculating the density from each subsample and taking an average.

However the calculation for taxa richness for a station is not equilvalent to taking the average taxa richness from the 3 subsamples. Station taxon richness should be calculated as follows:

Station Taxon Richness = all taxa observed at a station in all subsamples; not the average number of taxa of the 3 subsamples.

Replicate station

A replicate station is a specific, fixed sampling location within an area that can be recognized, re-sampled and defined quantitatively (e.g., latitude and longitude and a written description). For each habitat type, a number of replicate stations should be sampled, each resulting in a single composite sample preferably consisting of ≥ 3 benthic invertebrate field sub-samples. Stations located within the exposure area should be positioned so as to ensure exposure to the effluent plume. Additionally, sufficient physical separation should exist between the replicate stations to allow these to be considered statistical replicates.

The geographic extent of replicate stations for lakes, streams and rivers is recommended as follows:

> Lakes: The geographic extent of each replicate station should minimally be 10m x 10m and minimally be separated by at least 20 m.
>
> Rivers and streams: The geographic extent of each replicate station should encompass a longitudinal stretch of the river that includes one pool/riffle sequence. A river distance of 6 times the bankfull width should be adequate (Leopold et al. 1964, Newbury 1984, Leopold 1994) and allow a minimum separation of 3 times the bankfull width between stations of similar habitat. To ensure consistency in application for the EEM program, bankfull width is defined as in Newbury and Gaboury (1993) and methods are discussed in section 5.21. If it is not feasible to sample this length of river (e.g., large rivers or headwater streams with rapidly changing gradients), then an acceptable alternative approach would be to define the geographic extent of stations in a manner similar to that suggested for lakes (i.e., stations are re-visitable locations with predefined dimensions of at least 10 m x 10 m, with adequate separation).

Marine Coastal Environments: each of the replicate stations should be a defined location with re-visitable dimensions (e.g. 10m x 10m). Replicate stations may be spaced 50 m apart or more, depending on the size of the area. In some estuaries, a replicate station should encompass a longitudinal stretch which includes the major habitat to be sampled (e.g., a distance of 6X the bankfull width). If this length of river is not feasible for large estuaries, an alternative definition would be similar to that suggested for coastal areas.

Area

A **Reference Area**, as defined in Schedule 5 of the MMER, is "waters frequented by fish that is not exposed to effluent and that has a fish habitat that, as far as practicable, is most similar to that of the exposure area". Reference areas need not represent pristine conditions and may have anthropogenic impacts from other sources than the mine in question. For a detailed discussion of reference area selection see section 5.8.

The **Exposure Area**, as defined in Schedule 5 of the MMER, is "all fish habitat and waters frequented by fish that are exposed to effluent". For the benthic invertebrate community survey, the exposure area begins in the receiving environment as close as possible to the effluent discharge and continuing to a point where the benthic invertebrate community characteristics return to those found in reference areas. The exposure area may extend through several types of receiving environments (e.g., different stream orders, lakes or marshes, estuarine to marine, or intertidal to subtidal) and contain a variety of habitat types.

For **Initial** and **Periodic** Monitoring, the portion of the exposure area that should be studied and sampled for the benthic invertebrate community survey is termed the ***high exposure area***. The high exposure area is defined as the aquatic receiving environment that is most proximate to the effluent discharge, and that contains appropriate habitat type(s) with sufficient geographic area to accommodate the necessary number of replicate stations.

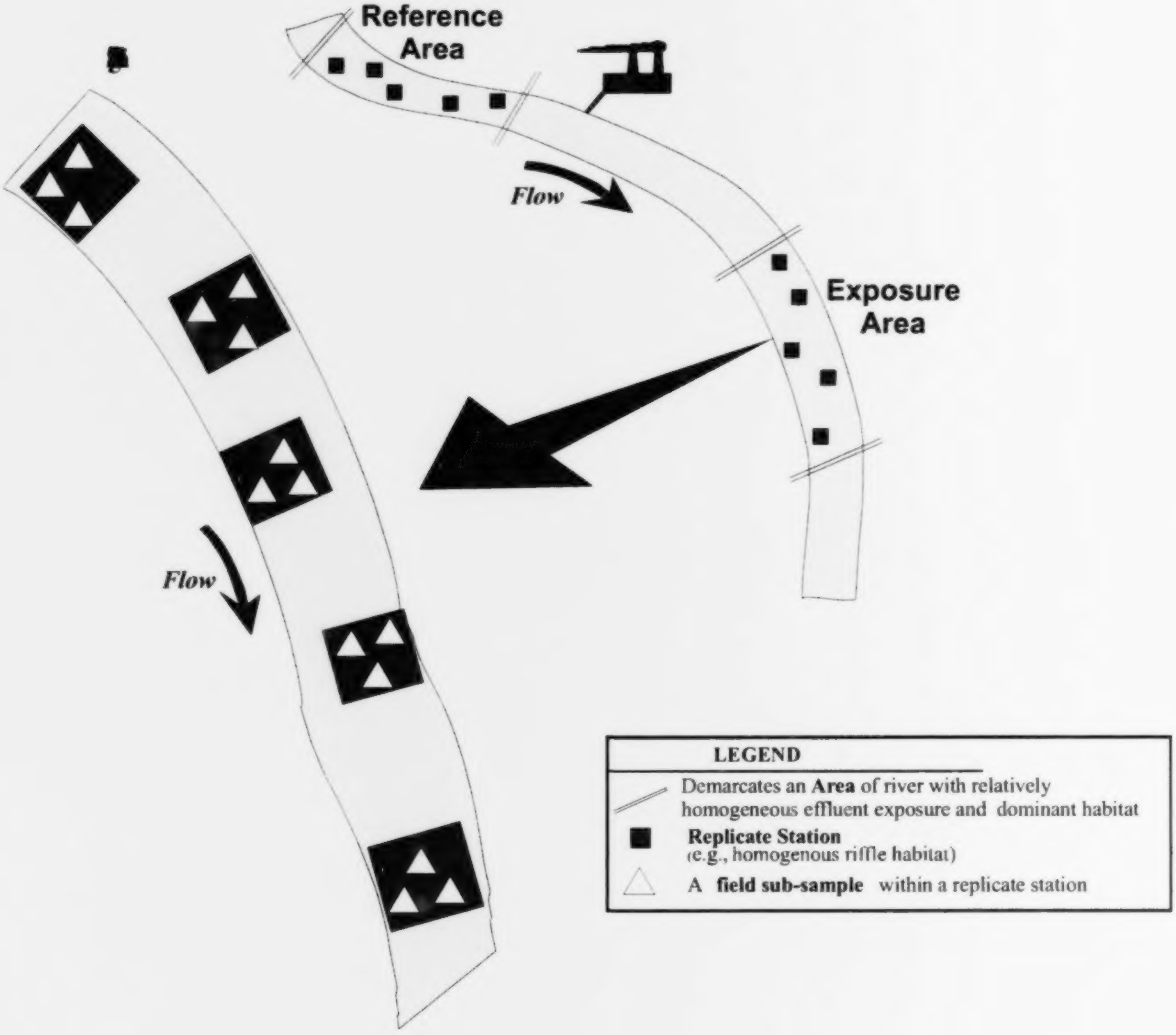
For **Focussed** Monitoring, the exposure area will extend along the effluent gradient so that additional areas with lower levels of effluent are included. The exposure area will extend geographically until a return to reference conditions.

Sufficient geographic coverage for a single study area is recommended for lakes, streams and rivers as follows:

> Lakes: The spatial extent of the study area should be at least 100m x 100m, and large enough to adequately accommodate the necessary number of replicate stations with sufficient separation.
>
> Rivers and streams: The spatial extent of the study area is defined in terms of stream or river morphology and should encompass a length of river, which is adequate to accommodate the necessary number of replicate stations with sufficient separation. The total length of river comprising an area would therefore be defined by the number of replicate stations multiplied by 6X the bankfull width, the river length, on average, in which one pool riffle sequence is expected to occur (Newbury 1984).

Estuary: For low salinity, relatively homogeneous estuaries, the area is defined in the same way as for freshwater rivers. For long, narrow marine regions such as narrow bays or fjords in which a control/impact type design is to be used, the area should be large enough to encompass the homogeneous habitat being sampled, as well as the defined exposure range. This will be at least 100m x 100m, and large enough to adequately accommodate the necessary number of replicate stations.

Figure 5-2 Examples of AREA, REPLICATE STATION and FIELD SUB-SAMPLE spatial scales for a basic Control/ Impact design.


Reference Area
Flow
Exposure Area
Flow
LEGEND
Demarcates an Area of river with relatively homogeneous effluent exposure and dominant habitat
■ Replicate Station (e.g., homogenous riffle habitat)
△ A field sub-sample within a replicate station

5.5.2 Reporting of Field Station Positions

The latitude and longitude in degree, minutes and seconds, as well as a written description of the sampling areas, is reported. The latitude and longitude co-ordinates can be obtained using a variety of methods. In some instances, the co-ordinates in conjunction with stream-wide distances (e.g. river kilometres), may be useful. Although, the positioning accuracy should be determined on a site-specific basis, the recommended specifications for positioning and recording sampling stations include the following:

- Co-ordinates should be expressed in geographic Universal Transverse Mercator (UTM) or Modified Transverse Mercator (MTM) using the North American Datum (NAD) of 1983 as the geodetic Datum;
- A local flat plane survey grid should be used only if it has been referenced to a minimum of three geographic UTM or MTM NAD 83 benchmarks;
- Bathymetry at effluent discharge points located on the Great Lakes should be referenced to depths below the International Great Lakes Datum (IGLD) using the most recent IGLD;
- If applicable, federal or provincial water level gauges should be monitored throughout the survey so that water depths can be accurately corrected to IGLD depths.

5.6 Sampling Program Designs

The design of the benthic invertebrate survey is site-specific and one of five benthic sampling program designs are recommended:

1) Control/Impact design (C/I)
2) Multiple Control/Impact design (MC/I)
3) Simple Gradient (SG) or Radial Gradient (RG) design
4) Multiple Gradient design (MG)
5) Reference Condition Approach (RCA)

These designs fall into three basic categories with different philosophical approaches. These are:

a) the Control/Impact or Multiple Control/Impact designs which are ANOVA type designs used to detect differences between discrete exposure and reference areas;
b) the Gradient (simple, radial or multiple) designs, which are intended to examine changes in community structure along a physical and/or effluent gradient which are better suited to regression analyses or ANCOVA; and
c) the multivariate approach of the Reference Condition, which compares potential "impaired" or test stations to a distribution of reference stations.

It should be noted that there may be some circumstances where ANOVA analyses are applicable to b) and c) above. Alternative monitoring methods (e.g., mesocosms) are also recommended (Table 5.2 and section 5.23) but must be scientifically defensible. A summary of the attributes, applicability and limitations of the five sampling designs is presented in Table 5.3 and described in more detail below.

The following descriptions apply primarily to the design of Initial and Periodic Monitoring programs, although special applications for determining the geographic extent of an effect for Focussed Monitoring are indicated, where applicable.

Control/Impact Design

The simplest study design for use in EEM is the Control/Impact (or reference-exposure) design (Green 1979). In rivers and in estuaries, this consists of no less than one reference area and a series of downstream exposure areas. For Initial and Periodic Monitoring, this should include, at a minimum, one high exposure area. (Figure 5.3A). Levels of exposure to mine effluent differ between the exposure and reference areas, but should be similar between the stations within each area. Habitat classes sampled should be consistent among areas and, with the exception of exposure level, these areas are to be as similar as possible in as many traits as possible (e.g., substrate, current velocity, water properties, environmental gradients, land use, etc.). This first study design employs ANOVA comparisons among areas and is recommended for simple, homogeneous rivers and streams without confounding upstream or near-site discharges from other sources.

The mine may propose modifications to this basic ANOVA approach providing that the modified design is scientifically defensible and addresses the appropriate monitoring questions. For example: if a reference area cannot be located upstream or in adjacent watersheds due to a confounding factor but a C/I design would otherwise be applicable, a modification of the C/I design may be appropriate. In this case, the design could be modified, so that the reference area is "downstream" instead of "upstream" (Figure 5.3B) of the point source. The downstream reference area would have to be outside of the area of significant exposure and meet the same reference area criteria as other designs.

This first study design is also recommended for simple, homogeneous estuaries or narrow inlets or bays without confounding upstream or near-site discharges from other sources, or where the ecologically relevant habitat occurs in spatially discrete but homogeneous patches (i.e., intermittent rocky outcroppings).

Focussed monitoring

The C/I design can be used to ascertain the geographic extent of an effect by first making use of rapid bioassessment protocols (US EPA 1999, Ellis and MacDonald, 1998), or other available information, to approximate how far the effect extends. Following this, a C/I monitoring program can be used that includes the high exposure area and targets additional exposure areas in localities where the effect is suspected to be dissipating (e.g., additional exposure areas located so as to bracket the suspected furthest reach of the effluent effects, together with the previous reference and exposure areas, see Figure 5.3C). ANOVA comparisons among areas can then be made to determine the geographic extent of an effect at a given significance level.

Multiple Control/Impact Design

Two of the major problems associated with the use of a single reference area are that it can be easily confounded by other factors and it is pseudo-replicated regarding the potential lack of independence among the stations in a single reference area (Hurlbert 1984). For example, a single reference area may differ from an exposure area with respect to several environmental variables in addition to the point-source effluent. In systems where an appropriate reference area is not available due to confounding factors, or it is determined, after review of historical information, that more reference areas are desirable, the multiple control/impact design should be used. Schematic diagrams of this design for application in either large rivers, lake and coastal-situated mines are presented in Figure 5.4. Sampling schemes should be devised so that the additional reference areas are located in adjacent watersheds or bays and that comparable habitat classes spanning the range of habitats found within the exposure area are selected.

The design philosophy of both the C/I and MC/I designs is that a specific difference in magnitude of effect between a series of areas is being examined. This lends itself to a classic ANOVA design with associated power analyses. These methods are statistically tractable and can provide indicators as to whether or not there is a biological effect from the mine effluent on the benthic invertebrate community. These designs assume that effluent exposure and habitat conditions are relatively homogeneous amongst all stations within a sampling area or that effluent exposure is within an acceptable range for a particular defined area.

Simple Gradient and Radial Gradient Designs

Simple and radial gradient designs (Figures 5.5 and 5.6) are suitable for situations where rapid effluent dilution precludes the selection of an exposure area that is comparatively homogeneous in terms of effluent concentration. As with the C/I design, gradient designs can be used in cases where no suitable reference areas are available upstream or in adjacent watersheds or bays. Gradient designs are also useful for determining how far along an effluent path effects are observed (i.e., geographical extent objective of Focussed Monitoring).

Philosophically, the gradient approach examines departures from expected (non-impacted) "patterns" of correlated biotic and environmental factors over spatial gradients. This is more suited to a regression type of analysis (or equivalent) in which replication (i.e., 5 stations within an area) is less appropriate than expending a similar effort to obtain accurate measurements of biotic and habitat variables over a sufficiently broad range of the gradient conditions. In the simplest case, a statistically significant effect would be declared if the slope of the regression of a response variable against distance from the effluent source is significantly different than 0. In this approach, a point source discharge is expected to have a "declining" effect away from the source. Therefore, it is not always feasible to make the judgment that there either "is" or "is not" an effect at a given station, but rather that there is a declining gradient of effects away from the point source. At a certain point along the gradient it is necessary to judge that this effect is no longer measurable or important. Therefore, in gradient designs, reference information is obtained from the stations furthest away from the effluent source.

A gradient does not necessarily imply straight lines, or the even spacing of stations within areas. The spacing of stations may be more or less continuous on a gradient away from the discharge, with less emphasis on distinctly different dilution zones and more on adequate geographic coverage, as compared to a C/I design. There are often no "blank" spaces between distinct sample areas, but rather a continuum of sampling stations along the gradient. However, if a change in effluent dilution within the receiving environment is abrupt, more sampling effort may be desirable over these stretches to accurately track rapid changes in mine effects. Wherever possible, the exposure gradient should be de-coupled or independent from any environmental gradients. For example, although a declining exposure gradient may fall along a path with varying depths, the two gradients may not necessarily be correlated. In cases where the exposure gradient is correlated with a co-occurring environmental gradient, a multiple gradient design may be more appropriate (see next section).

Simple gradient designs are particularly appropriate for narrow water bodies, such as rivers and streams. In wider water bodies, such as lakes or open coastal areas, a radial gradient design can be more appropriate. Sampling is conducted away from the effluent source along several gradient transects (Figure 5.6). As in the multiple C/I approach, the use of a radial gradient design will provide a larger number of reference sites. Furthermore, a broader geographic area will be sampled which can be important in non-homogeneous, open lakeshore or marine areas, which often have complex current and circulation patterns or a variety of equally important major habitat classes or gradients.

Wherever possible, the exposure gradient should be de-coupled or independent from any environmental gradients. For example, a declining exposure gradient may fall along a path with varying depths, but a simple or radial gradient approach may still be feasible if the exposure and depth gradients are not correlated and the differences in depth are not so great as to obscure any effluent effects. Alternatively, a multivariate approach may be necessary to remove the confounding influence of varying depth.

For radial gradients, a comparison of regression patterns for each gradient (e.g., regressions of faunal abundance versus distance from the outfall) may help to illuminate the direction and extent of effects. Alternatively, all data from all gradients can be included in one regression, if the comparison is between biotic and physical factors unrelated to geographic or natural habitat factors. If sufficient sampling is done (e.g., radial gradients), it may be possible to pick and choose unconfounded replicate stations (e.g., homogeneous habitat conditions) to regress a biotic versus a mine-related variable.

Focussed Monitoring

Due to the layout of sampling stations, gradient designs are particularly well suited for determining the geographical extent of an effect. The simplest design for Focussed Monitoring would be to allocate sampling stations along a gradient from more to less exposed, ensuring that the most distant stations are located well beyond the likely extent of effects. The geographic extent of effects could then be determined graphically by plotting the response variable(s) against distance from the mine and inspecting the data for an inflection point where the response variable asymptotes to the reference condition. Data from sampling stations arrayed in this manner could also be used, together with measured physicochemical data, in a multivariate analysis (e.g., ordination or clustering) used to identify which distant stations tend to group with reference stations and which tend to group with clearly impacted stations. Both of these approaches (graphical plotting and multivariate analysis) look for patterns in the data to determine the approximate extent of an effect; that is, they do not entail hypothesis testing and, therefore, power analysis would not be applicable in these cases (in contrast to the C/I approach to Focussed Monitoring given above).

It is also possible to design a hypothesis-testing gradient program for examining the geographic extent of an effect. This would entail using field-subsamples as replicates (treating stations as areas) and making station-by-station ANOVA comparisons along a gradient to determine where an effect disappears at a given significance level. This latter approach might, however, require extensive sampling effort, depending upon the number of stations along the gradient and the required (by power analysis) number of field sub-samples per station.

Multiple Gradient Design

In some cases, it may also be useful to compare "affected" versus "reference" gradients. This would be the case when a co-occurring environmental gradient confounds an effluent gradient in the exposure area. By using a multiple gradient design (Figure 5.7), it is possible to make statistical comparisons of the exposure area gradient to a similar environmental gradient in a reference area(s). The reference gradient(s) should be as similar as possible in depth and habitat to the exposure gradient. Potential effluent impacts would be tested for by using ANCOVA to factor out the influence of the co-occurring environmental gradient.

Reference Condition Approach

The fundamental concept of the Reference Condition Approach (Figure 5.8) is to establish a database of sites that represents unimpaired conditions (reference stations) at which biological and environmental attributes are measured. This database is used to develop predictive models that match a set of environmental variables to biological conditions. These predictive models then allow a set of environmental measurements to be made at a new station and used in the model to predict the station's expected biological condition (i.e., the biotic conditions of the group of reference stations with similar environmental attributes). A comparison of the actual biological condition at the new station with the predicted conditions allows an assessment of the condition of the new station to be made.

The RCA can reduce the need to find nearby comparable reference sites when studying an impacted system, which can be a problem in some traditional approaches. Rather than identifying and sampling upstream reference sites in a river system or next-bay-over reference sites in a lake, the RCA uses a set of biologically equivalent reference sites selected from an existing database to evaluate an exposure site. Provided that it is kept up-to-date, the reference condition database can be used over a number of EEM phases.

The reference condition database is established by an initial standardised sampling program at a wide variety of geographic scales. The same benthic invertebrate community sampling protocol is used in as many ecoregions and stream orders or lakes as are available in a catchment. A number of environmental variables are measured in conjunction with invertebrate sampling. The data are then subjected to a 3-step multivariate analysis in which:

1) a number of invertebrate groups are formed based on similarity of community structure,
2) biological data are correlated with environmental attributes, and an optimal set of environmental variables is identified that can be used to predict group membership, and
3) the biological condition of test (exposure) stations is assessed by using the optimal set of environmental variables to predict group membership. How the test station fits, relative to the group to which it is predicted to belong, establishes whether, and to what degree, the station is different from the reference group. Assessment can be made by either the use of the community descriptors described above, by determining if the site is within the prescribed range of variation observed at reference sites (2 SDs), or by the use of ordination methods and determining if the reference site is within the 95% probability ellipse of the matched reference sites.

Depending on the timing and locations of the sampling program, it may also be possible to use the resulting database to make ANOVA comparisons between reference and exposure areas.

Once the reference database is established, the RCA can be used as a rapid bioassessment method and can be used to deal with national and local-scale issues using the same database and software. Due to the intensive initial sampling effort required, the RCA would not be considered a practical approach for use by a single mine in a remote location if a reference database is not already available; however, it may be applicable in areas where multiple industries (including different EEM industrial sectors) are located. In this case it may be practicable and cost effective for multiple users to collaborate in the development of the reference database.

5.7 Alternative Monitoring Methods

If there are confounding factors and they cannot be resolved through modification of the study design, the mine may choose a scientifically defensible alternative monitoring method (Table 5.2). Currently the only recommended alternative method for the benthic invertebrate component is the application of mesocosms to conduct on-site community bioassays, however, other scientifically defensible alternative monitoring methods, that can determine if the mine effluent is having an effect on the benthic invertebrate community, may be proposed by the mine.

Mesocosms provide a means for the manipulation of aquatic populations and communities under controlled effluent exposure conditions with sufficient replication, precise control over relevant variables under environmentally relevant field conditions (i.e., temperature, water quality, photoperiod). Thus, properly designed mesocosm studies will allow for the impact of the mine effluent to be examined independently from confounding factors present at a particular mine site. Although, mesocosms range widely in design (Shriner and Gregory 1984, Kosinski 1989 , Beyers an Odum 1993, Cooper and Barmuta 1993, Guckert 1993, Graney et al. 1994, Pontasch 1995, Culp et al. 1996), recent examples of the use of mesocosms for the assessment of effects in freshwater and marine receiving environments within Canada include Culp et al. (1996, 2000), Lowell et al. (1995), and Dube et al. (1998). Mesocosms are also useful as an Investigation of Cause Tool (section 5.23.1) and their applicability and methodology is described in detail in a guidance document for the use of mesocosms in EEM programs (Chapter 12, Section 12.1).

Table 5-3: Recommended benthic invertebrate community sampling program designs.

Note: Multivariate analyses can be performed on data collected using any of the above designs to look for patterns (i.e., not hypothesis tests) that may be useful for highlighting potential areas of concern.

Design Type	Receiving Environment	Reference / "control" Areas	"Impact" Areas	Statistics
Control/Impact (C/I) *Figure 5.3*	freshwater rivers or lakes, homogeneous or low salinity estuaries	a single reference area, upstream of mine discharge	High Exposure Area for Initial Monitoring (additional exposure areas are added in Focussed Monitoring)	ANOVA
Multiple - Control/Impact (MC/I) *Figure 5.4:A,B&C*	Freshwater, rivers or lakes with geographicly homogenous lake shores, homogenous estuaries and coastal zones	Multiple reference areas in the same or environmentally similar adjacent watersheds or bays	High Exposure Area for Initial Monitoring (additional exposure areas are added in Focussed Monitoring)	ANOVA
Simple Gradient (SG) *Figure 5.5:A&B*	freshwater rivers or geographicly restricted lakes, non-homogeneous narrow estuaries or geographicly restricted marine bays, inlets or fjords	a series of reference stations with no or low effluent levels, situated towards the end of a declining gradient of mine effluent	single gradient through declining levels of effluent in the receiving environment	Regression/ ANCOVA
Radial Gradient (RG) *Figure 5.6*	freshwater lakes, non-homogeneous open marine bays and coastal areas	as above, but situated toward the end of several radially oriented transects	as above, but repeated in a radially oriented design	as above
Multiple Gradient (MG) *Figure 5.7:A&B*	Freshwater lakes or rivers Non-homogeneous open marine bays and coastal areas	a series of reference stations with no effluent situated on a transect along the same kind of environmental gradient observed in the exposure area	gradient through declining levels of effluent and a co-occurring environmental gradient in the receiving environment	ANCOVA, with reference and exposure transects considered as treatment groups
Reference Condition Approach (RCA) *Figure 5.8*	Freshwater rivers or lakes, particularly for cooperative investigations or where there is an existing reference database	multiple series of reference stations with no or low effluent levels situated in similar drainage basins within the same ecoregion	series of stations within the exposure area which are tested individually against the reference station distribution	Multivariate /ANOVA (if possible)

Figure 5-3a: Control/Impact designs for simple freshwater rivers and streams or for homogeneous estuarine habitat.

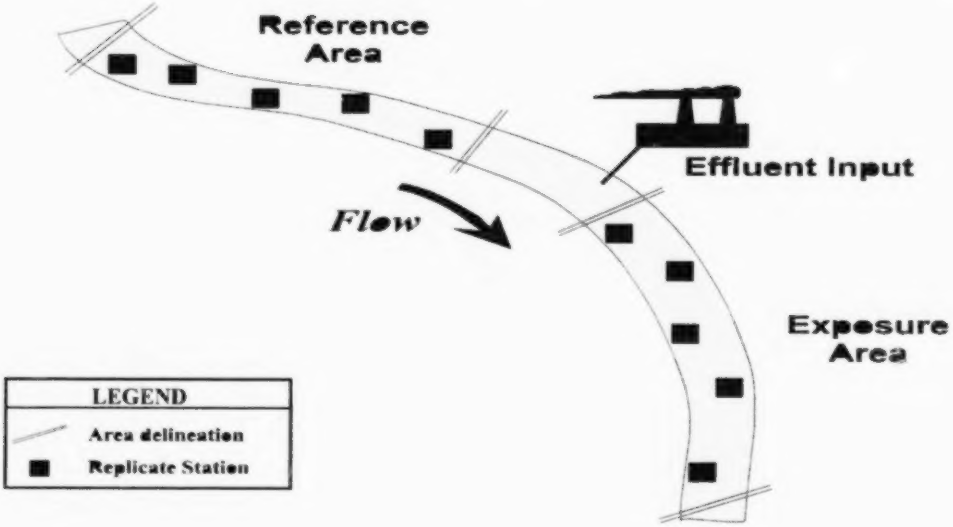

Reference Area
Effluent Input
Flow
Exposure Area
LEGEND
Area delineation
Replicate Station

Figure 5-3b: Modified C/I Design with downstream reference area for stream, river or estuaries

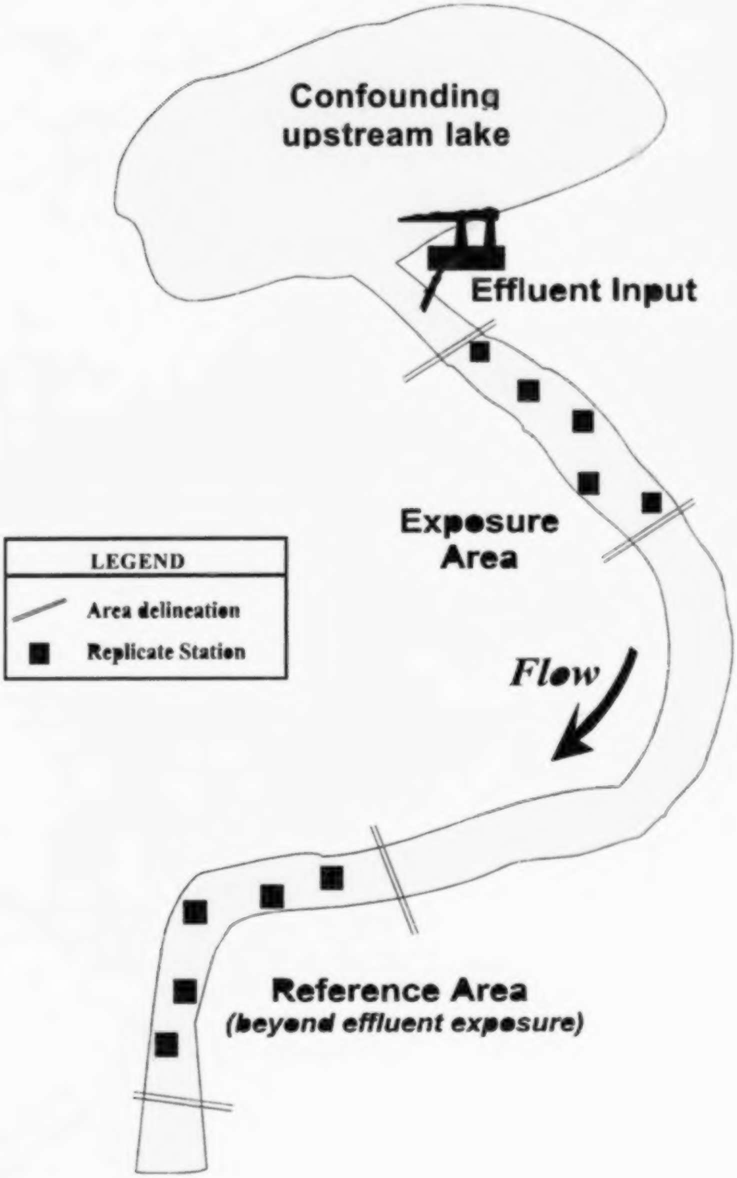

Confounding upstream lake
Effluent Input
Exposure Area
LEGEND
Area delineation
Replicate Station
Flow
Reference Area
(beyond effluent exposure)

Figure 5-3c: Focussed Monitoring design example (currently under review)

Figure 5-4a&b: Multiple control/impact designs for freshwater rivers and streams with two reference areas.

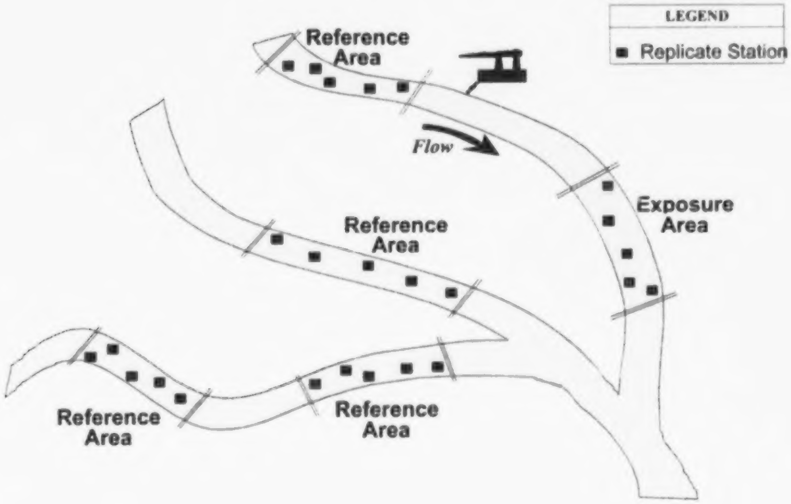

LEGEND
■ Replicate Station
Reference Area
Flow
Exposure Area
Reference Area
Reference Area
Reference Area

Figure 5-4c: Freshwater, geographically homogeneous lakes and marine bays or inlets (currently under review)

Figure 5-5A: Simple gradient designs for freshwater rivers or streams, and estuaries

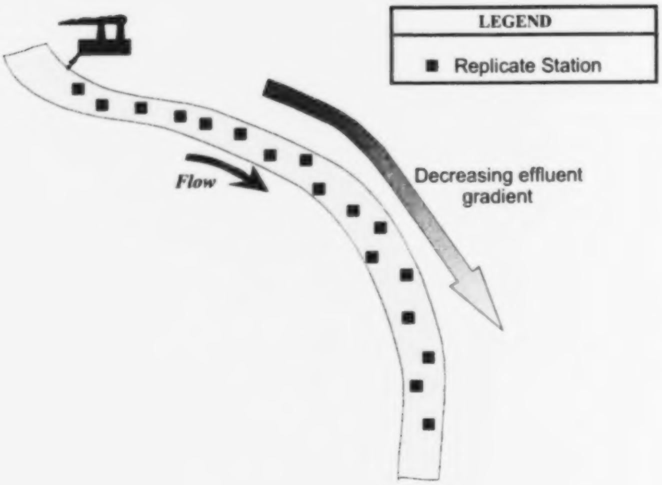

LEGEND
■ Replicate Station
Flow
Decreasing effluent gradient

Figure 5-5b: Simple gradient designs for lake or coastal mines situated in narrow bays or fjords.

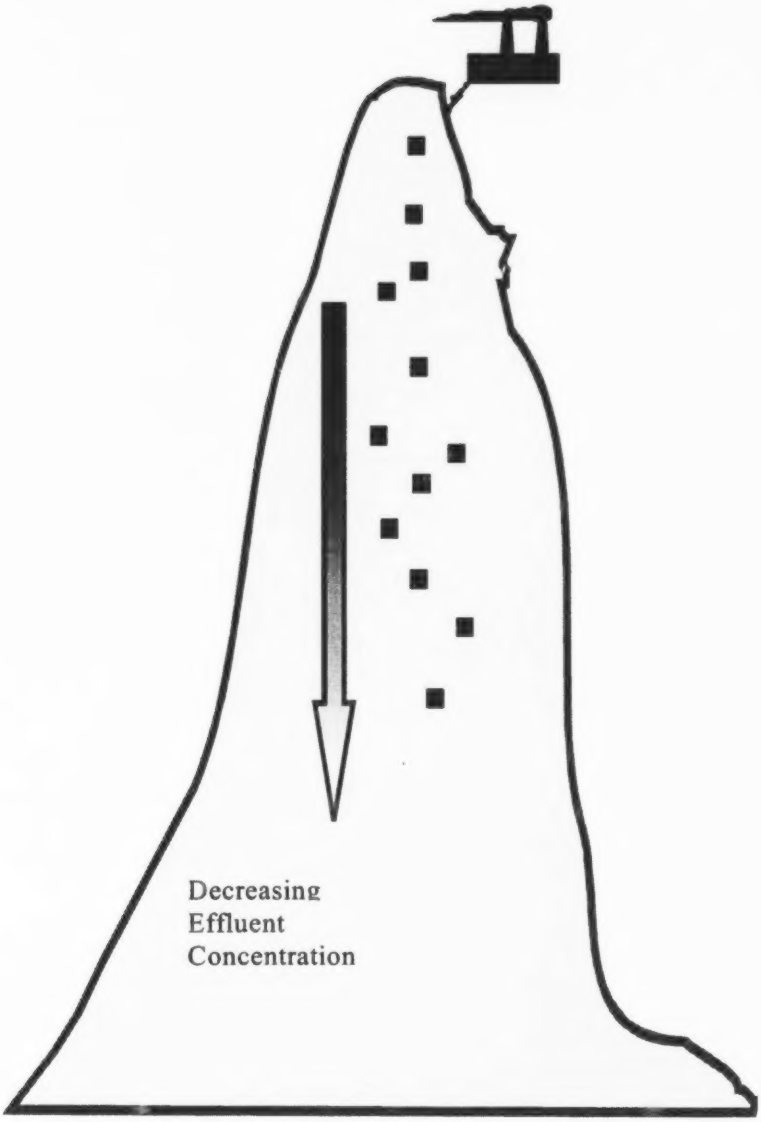

Decreasing
Effluent
Concentration

Figure 5-6: Radial gradient design for lake or coastal situations

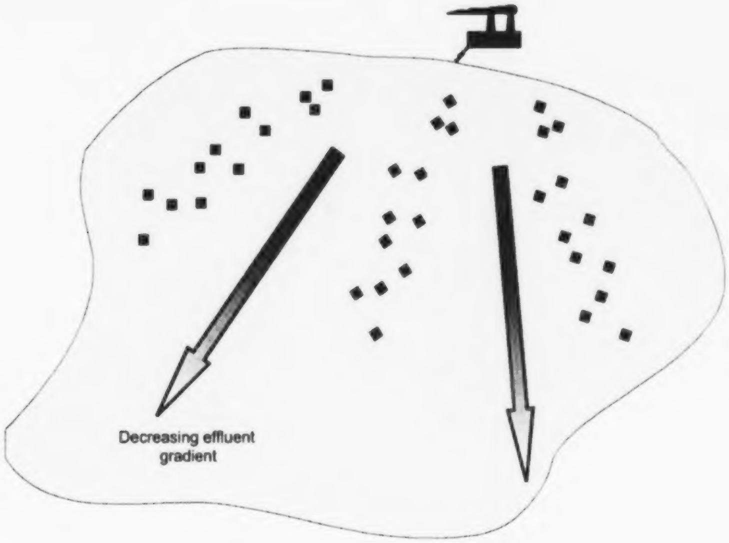
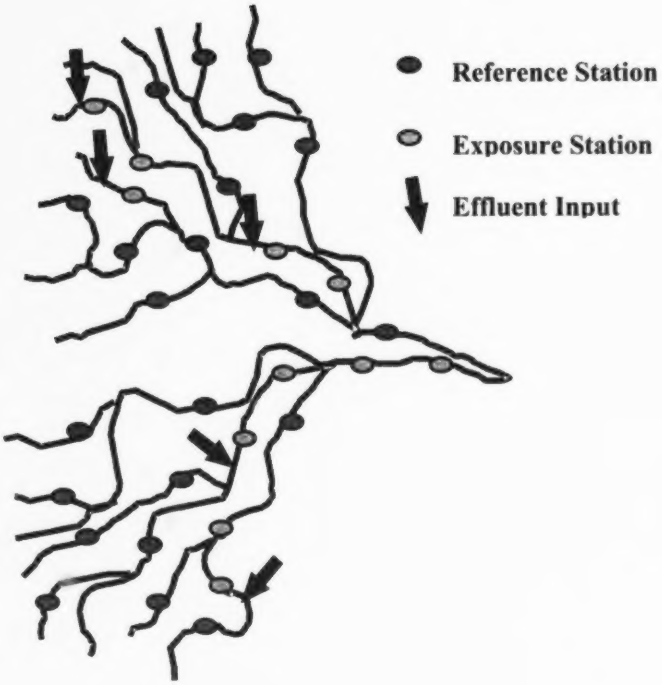

Decreasing effluent gradient

Figure 5-7a&b: Multiple Gradient Designs (currently under review)

Figure 5-8: Reference Condition approach


● Reference Station
○ Exposure Station
Effluent Input

5.8 Selection of Reference and Exposure Areas

As a first step in determining the most appropriate study design for any monitoring program, it is necessary to locate suitable reference areas. When environmental and biological data obtained from reference areas are compared to exposure areas, they can detect impairment of aquatic life (Yoder 1991), diagnose stressors (Hughes et al. 1994), provide data on temporal and geographic trends (Yoder 1989) and provide data for water resource summaries for government agencies (OEPA 1990). Identification of appropriate reference areas is essential for the EEM program and should initially be done during Site Characterization, although re-evaluation or addition of reference areas may occur during other EEM phases. Careful consideration should be taken, not only for the consistent application of the selection criteria for reference areas, but also in determining how many reference areas or stations will be examined for particular study designs.

5.8.1 General guidance for Reference Area Selection

A reference area should have no effluent exposure from the mine in question and natural habitat features similar to those of exposure areas. However, reference areas do not necessarily represent pristine (or pre-European settlement) conditions, but areas in which anthropogenic impacts, unrelated to the mine effluent, are similar to exposure areas (Simon 1991, Omernik 1995). For example, Omernik (1995) identifies references areas in drainage basins that lack municipal and industrial influence even within the polluted Huron-Erie corridor of the Great Lakes. A second example can be found on the prairies where minimally impacted reference areas may be located in drainage basins that lack point source effluents but, unavoidably, have a high degree of agricultural activity. Although, the characteristics of reference areas will differ across the country, the approach to defining and identifying reference areas should be consistent.

Because reference areas will vary among different landscapes, approaches have been developed which schematically classify river and lake ecosystems. One of the objectives of these classification schemes is to predict the aquatic biotic assemblages, which are likely to occur in a particular location (Corkum 1989, 1992, Hughes 1995, Maxwell et al. 1995, Omernik 1995). A hierarchical classification system is a way of simplifying sampling procedures and management strategies by organizing a variable landscape by discriminating among features at several scales of resolution (Conquest et al. 1994). Two recently proposed classification schemes for characterizing aquatic habitats are compared and summarized by Bisson and Montgomery (1996). Both systems are based on hierarchies of topographic and fluvial characteristics and both employ descriptors that are measurable and ecologically relevant. The recommended framework for classifying aquatic features is the classification system developed by the U.S. Fish and Wildlife Service (FWS), "Classification of Wetlands and Deepwater Habitats of the United States" (Cowardin et al. 1979, Busch and Sly 1992) and these methods are discussed more fully in Chapter 3.

5.8.2 Habitat-Specific Allocation of Reference and Exposure Areas

The following specific points should be considered during the selection of reference and exposure areas and/or stations (for gradient or RCA designs) for a EEM benthic invertebrate program.

- Identification of the exposure area and its habitat features should precede the selection of reference areas because reference areas must match the physical and chemical habitat features of the exposure area (other than those expected to change due to the effluent).
- Selection of an appropriate classification scheme is important. The recommended framework for classifying aquatic features is the classification system developed by the U.S. Fish and Wildlife Service (FWS), "Classification of Wetlands and Deepwater Habitats of the United States" (Cowardin et al. 1979, Busch and Sly 1992). Other examples of hierarchical classification schemes include those of Maxwell et al. (1995), Conquest *et.al.* (1994), and Meador et al. (1993).

- In order to determine if a reference area is physically suitable or not, referral to detailed classification systems as outlined for marine shorelines to deep coastal areas include Frith et al. 1993, Booth et al. 1996, Robinson and Levings (1995), Hay et al. (1996) and Robinson et al. (1996). Specifically, estuarine classification has been reviewed by Matthews (1993), Scott and Jones (1995), Finlayson and van der Valk (1995) and Levings and Thom (1994). In the US, the most widely used system is that of Cowardin et al. (1979), Cowardin and Golet (1995) with expansions proposed by other authors. Other sources are cited in Watson (draft report).

For Rivers

- The size of the drainage basin selected is based on stream order. For example, if a mine site is located on a second order stream, drainage basin area is delineated at the point the stream becomes third order (i.e. at the junction of two second order streams).
- If there are no upstream inputs or confounding factors, reference area(s) can be within the drainage basin and upstream of the mine.
- If confounding factors, such as non-point or point source inputs, occur upstream of the effluent, reference area(s) can be selected in a nearby drainage basins with comparable habitat features (Figure 5.4).
- If physical disturbance of the river valley is associated with the mine, effluent effects may be confounded by the disturbance. Accordingly, reference areas should be selected, if possible to match the physical disturbance.

- The following features should be similar between reference and exposure areas:

  - ecoregion, drainage basin area, stream order, bankfull width, channel gradient, channel pattern, habitat types, water depth, water velocity substratum composition, riparian vegetation, shoreline structure and land use.

For Lakes

- If tracer studies indicate that effluent effects are local and restricted, select reference areas within the lake in which the mine discharge occurs. These reference areas should occur in separate, but comparable bays or basins of the lake.
- If tracer studies indicate that the identified effluent is dispersed throughout the lake, select reference area(s) in the nearest comparable lake within the same or adjacent drainage basin.
- If non-point or other point source inputs occur elsewhere on a lake, select reference area(s) in the nearest comparable lakes within the same or adjacent drainage basin.
- If the mine effluent is associated with physical disturbance in the area, effluent effects may be confounded by the disturbance. Accordingly, physically matched reference areas should be selected, if possible.
- The following features should be similar between reference and exposure areas
  - ecoregion, geological origin, drainage basin area, morphometry, slope from shoreline, habitat types, substratum composition, riparian vegetation, shoreline structure and land use

For Marine Environments

- The reference area should be within the same water body and hydrographic current or tidal regime as the exposure area. In other words, the closer the reference area is to the exposure area, the better. Benthic invertebrate communities are considerably higher in species richness, have more complex trophic relationships, faunal size ranges and reproductive strategies in marine than in freshwater. This is simply because there is a much higher diversity of types of organisms in the marine benthic invertebrates than in freshwater. Because of this complexity, and the multitude of interactions between species in marine benthic invertebrates, small shifts in physical or chemical conditions can dramatically alter the overall benthic faunal community. Add to this the effect of increasing variation in chance larval settlement with increasing geographic distance (geographic "drift" in community structure) and physical barriers in complex coastlines, and it is very rare to find similar invertebrate communities from one bay or fjord to the next, or to be able to predict specific benthic community structure based on sediment factors (for recent review on invertebrate sediment interactions in the marine see Snelgrove and Butman 1994). In order to have some confidence that the "natural" benthic community is similar enough from one coastal area to the next, there must be sufficient water exchange between them.

This is more likely in open coastal areas than in isolated bays and fjords.

- Reference areas which are not in the same water body or hydrographic regime may only be suitable for comparisons of summary characters such as shifts in abundance or species richness. If the habitat conditions are similar enough to the exposure area, it may also be possible to compare larger-scale biotic factors, such as the presence of characteristic, long-lived depth/substrate specific taxa described by Thorson (1957) as "parallel communities".
- The reference area should have very similar habitat type, shoreline structure (steep, mountainous, delta, marsh, etc.), bottom topography (sills, sandbars, exposure to open oceanic influences, etc.), substrate type (particle size, sorting, natural chemistry), depth properties, current regimes, physical water properties, nutrient regimes, confounding inputs and drainage characteristics as the mine survey area.
- Some special considerations are important for determining the suitability of reference areas for marine and estuarine mines. Physical factors in the estuarine/marine environment which tend to be more complex than in freshwater are: salinity (including seasonal freshwater influence), tides (and tidal currents), and sediment sulfides. Other important physical factors include ice-scour or build-up, freezing, water column stagnation due to large summer freshwater run-off, resuspension due to surface freezing in winter (see Ellis et al. 1994), dams or log-booms; extraordinary siltation or clogging from logging, periodic flooding, etc.
- In addition to the above important characteristics , the following specific points should be similar between reference and exposure areas:

⇒ Intertidal areas: shoreline slope, wave exposure, light and tidal exposures, shoreline vegetation, and encrusting fauna (although the latter may be part of the benthic taxa being monitored for a response to mine effluent).

⇒ Subtidal areas: seasonal water column stability and bottom oxygen depletion (stagnation) .

5.8.3 Reference and Exposure Area Consideration for the EEM program phases

The allocation of reference and exposure areas is dependent on the site-specific study design (section 5.6), and the phase of the EEM program.

For **Initial Monitoring**, the objective is to determine whether there is an effect on the benthic invertebrate community in the high exposure area where an effect is more likely to occur. This spatial limitation is designed to concentrate sampling effort in a cost-effect manner. The study design and allocation of reference and exposure areas should be based on this objective.

For **Periodic Monitoring**, the objectives are to confirm results, detect changes and allow for trend monitoring data. As these objectives are similar in geographic scale to Initial

Monitoring the selection criteria for reference and exposure areas should be the same. However, as with any on-going monitoring program, the appropriateness of reference and exposure area selection should be re-evaluated as additional information is gained.

For **Focussed Monitoring** there is a further objective, to determine the geographic extent of previously identified effects. Thus, sampling should be conducted at exposure areas located at further geographical distances away from the mine effluent discharge, until a return to reference conditions. The physical allocation of multiple exposure areas and stations is dependent on the study design. If a confounding factor is encountered before the reference condition is reached and this factor cannot be resolved by modifications in study design (Table 5.3). The exposure area may be defined to extend only as far as to where the confounding factor is encountered. Alternative, cost-effective study designs or methods may be applicable (section 5.23, Table 5.2).

In addition, as part of the review of monitoring information, reference areas sampled in previous monitoring should be re-evaluated to assess whether they are adequate for the **Focussed Monitoring** program. The new geographic extent of Focussed Monitoring may naturally include additional habitats, such as higher order streams or lakes. If these new habitats were not represented in the reference area(s) that were used for previous monitoring, a re-assessment of the adequacy of these reference areas will be necessary. The addition of reference areas should also be considered to allow a more balanced design between the number of reference and exposure areas.

If a RCA study design was used during previous monitoring, additional reference areas may not be necessary (assuming they adequately represent the habitat types), but it is recommended that a subset be re-sampled to examine the effects of natural temporal variation.

5.9 Selection of Ecologically Relevant Habitats

5.9.1 General Guidance for Habitat Selection

The most ecologically relevant habitat(s) should be sampled within the exposure area(s) and similar habitat(s) will be located and sampled within the reference area(s). The selection of the appropriate habitat type(s) requires consideration of the following questions:

- Which habitat type is present in the highest proportion in the exposure area?
- Which habitat, in the absence of human influences, supports the richest assemblages of invertebrates (benthic invertebrate diversity) within the study area?

- In which habitat are the invertebrates most likely to be exposed to sediment or water borne contaminants for extended periods of time?
- Is historical information available for a particular habitat?

To identify the ecologically relevant habitat, the first consideration is to sample the habitat that accounts for the greatest proportion of the exposure area. However, other factors can override the importance of geographically dominant habitat including the ecological relevance of sampling highly sensitive and diverse habitats even if they comprise a lower proportion of the study basin. In streams, riffles can support a diverse assemblage of species that display a wide range of sensitivities to water-quality changes. Therefore, the community in this habitat has the potential for greater change than less species rich communities. In contrast, the fauna of depositional areas, which are generally less rich taxonomically, are of interest during biomonitoring exercises because they may be directly exposed to concentrations of sediment borne contaminants for longer periods. Consequently, communities in depositional areas may respond to contaminants differently than the more sensitive but less exposed riffle habitats. For additional guidance regarding stream habitat selection refer to Cuffney et al. (1993), Plafkin et al. (1989) and Meador et al. (1993).

5.9.2 Habitat Considerations for the EEM program phases

Decisions on which habitat to sample should be a site-specific consideration. Decisions regarding the sampling of greater than one ecologically relevant habitat during the same survey are dependent upon the phase of the EEM program.

For **Initial Monitoring**, the objective is to determine whether there is an effect on the benthic invertebrate community and the habitat most likely to exhibit these effects should be sampled. If more than one habitat is determined to be ecologically relevant, the effort could be expended during Initial Monitoring to sample all ecologically relevant habitats. This may reduce the potential of missing an effect in a sensitive habitat and/or the necessity of expanding the survey to additional habitats during future monitoring efforts. If questions regarding confirmation or magnitude and extent can be addressed by additional sampling during the same field trip it may be cost effective to do so.

For **Periodic Monitoring**, one of the objectives is confirmation of previous results. The appropriate selection of ecologically relevant habitats should be confirmed during this monitoring phase and could also be used to address questions such as:

- Are there ecologically relevant habitat types that were not previously sampled but may be more sensitive to the site-specific environmental and effluent conditions?

For both **Initial and Periodic Monitoring**, the sampling of additional ecologically relevant habitat types should not be at the expense of reduced sampling effort in the primary habitat of interest. For most biomonitoring studies, the sampling of a single habitat is intended to reduced the variability which is inherent in sampling natural substrates but which would be even greater if the same level of effort is spread over a range of habitat types.

As with Initial and Periodic Monitoring, for **Focussed Monitoring** the most ecologically relevant habitat(s) will be sampled within the exposure areas and similar habitat(s) will be located and sampled within the reference area(s). The decision regarding number and type of habitats to sample is made based on a review of the previous monitoring results, site specific considerations and the objectives of Focussed Monitoring. For example, to determine the geographical extent of an observed effect, additional habitats such as higher order streams or lakes may become important. On the other hand, if during previous monitoring a number of habitat types were sampled but one particular habitat appeared to show responses and the others did not, this habitat type could be targeted during Focussed Monitoring.

For Marine/Estuarine Habitats

A decision needs to be made to sample either depositional or erosional habitats in the estuarine/marine receiving environment. In addition, decisions about sampling intertidal versus subtidal substrates for estuarine/marine mines will depend on which is the appropriate receiving environment, and on the feasibility of obtaining useful samples. In marine/estuarine habitat, the selection of the appropriate habitat type(s) therefore requires consideration of the following questions:

- What habitats are feasible to sample?

The habitat which is most common geographically and most likely to be affected by the effluent should be selected. However, selection of major habitats is partially related to viability of sampling. For example, if the major habitat is a vertical rock cliff to a depth of 300 m immediately outside the effluent discharge, this is rarely feasible for benthic invertebrate sampling without extraordinary equipment. As another extreme example, if the major habitat is intertidal, but consists of a steep rock cliff with heavy wave and wind exposure or ice build-up, sampling may not be possible. When multiple habitats are available and appropriate, some choices must be made. In some cases, more than one habitat may have to be sampled (radial gradient or similar design). Where there is a choice, sampling of soft substrates is preferred because methods are generally more quantitative.

- What is the habitat that is most biologically active or "important"?

When the subtidal environment most exposed to mine effluent consists of both solid and unconsolidated sediments, then either both substrates need to be sampled or a decision on which to sample must be taken. All other factors being equal, the unconsolidated sediment is the most efficient to quantitatively sample. However, when it is obvious that a coarse sand substrate is almost devoid of macrofauna within the top 10 cm (or the depth of penetration of a sampling device), whereas the nearby rocky reef is extremely rich and an obvious haven for many fish, it is the most "active". Similarly, if there is an important fisheries resource in one major habitat type which is directly exposed to mine effluent, it may be considered the most biologically important.

- Can the ecologically relevant habitat be "classified" according to recognized physical type and characteristic species?

Habitat classification systems have been discussed and reviewed by many researchers and can be useful for comparing "expected" biotic factors with actual biotic factors present in the mine vicinity. Some recent marine examples include a comprehensive draft document presented to DFO for delineating the Strait of Georgia on the west coast of Canada and northwest US (Watson, draft). Some relevant references for marine classification for worldwide shorelines to deep coastal areas include Frith et al. 1993a, Booth et al. 1996, Robinson and Levings (1995), Hay et al. (1996) and Robinson et al. (1996). Specifically, estuarine classification has been reviewed by Matthews (1993), Scott and Jones (1995), Finlayson and van der Valk (1995) and Levings and Thom (1994). In the US, the most widely used system is that of Cowardin et al.(1979) and Cowardin and Golet (1995), with expansions proposed by other authors.

- Is the effluent discharge depth and/or buoyancy most likely to affect the intertidal or subtidal regions?

If the effluent discharge is and remains mainly intertidal, then this should be the targeted habitat. However, if the effluent affects both intertidal and subtidal habitats, then the subtidal is the preferred habitat, because this area is most likely to show impacts in fish. If suitable, both habitats may be sampled. This question must also take into consideration seasonal water column stability changes, which can affect intertidal areas. For example, at the Black Angel Mine in Greenland, surface sea-ice in winter created a homogeneous water column causing re-suspension and shallow deposition of settled tailings (Ellis et al. 1994).

- What habitat type is present in the highest proportion?

In many cases, coastal shorelines will be mixed silt, sand, gravel and rock substrates. In bays near freshwater discharge, there tends to be accumulations of soft substrate such as sand or silt. Estuarine mines will usually have dominantly soft substrates from river-borne material. If there are similar percentages of both depositional and erosional habitats, the preferred habitat to sample is depositional, because this type will accumulate the discharged material from mine effluents and is more likely to suffer deleterious effects from organic enrichment. Erosional substrates tend to be kept "clean" by high current action or wave or ice scouring.

However, if the percentage of solid substrate habitat is much greater than soft substrates, or if a previously "clean", rocky shoreline has begun to accumulate sediment related to mine discharge, then this may be the preferred habitat to sample.

- Are there confounding factors which may affect benthic communities?

Benthic communities in naturally or anthropogenically confounded sampling areas are problematic to use for interpreting effects from mines. Obviously they should be avoided. For example, in situations where solid and unconsolidated substrates are present, only one of these may be outside the influence of the confounding factors. One source of confounding factors that is particularly important in arctic areas is the seasonal or year-round effects of freezing or ice-scour, particularly in the intertidal or estuarine areas, which may seriously disrupt surficial communities.

- What is the environment affected by subtidal discharges?

Obviously, the environment most exposed to effluent should be the targeted sampling area and will also determine the type of sampling design used. In an estuary, if the discharge occurs at the surface where there is a strong and permanent surface freshwater layer with little intrusion of saltwater at high tide at depth, then the habitat to be sampled is downstream from the mine However, where there is a strong tidal intrusion, sampling will have to go upstream and downstream. There are numerous other factors of this type to consider, all of which require detailed site-specific information about the habitats and the pattern of effluent dispersal.

In summary, if there is a choice of habitats to sample, it is recommended first that subtidal habitats be sampled because they tend to have higher diversity and less patchiness in fauna than the intertidal, due to less extreme or harsh habitat conditions. This is particularly true in arctic regions where extreme winter conditions may eliminate most of longer-lived fauna that tend to more clearly integrate the effects of contaminants. Secondly, if there are a variety of suitable habitats, depositional habitats should be chosen, particularly for subtidal areas, because the methodology allows easier and more quantitative sampling procedures. In addition, depositional areas tend to accumulate contaminants over time, whereas erosional areas may not.

5.10 Selection of Ecologically Relevant Sampling Seasons

5.10.1 General Guidance for Sampling Season Selection

All benthic invertebrate community surveys should be completed during the most ecologically relevant season. For mines which do not discharge continuously, sampling should occur during a period of effluent discharge and after the receiving environment has been exposed to the effluent for a sufficient period for which effects would reasonably be expected to occur (i.e., generally within 3- 6 months).

The preferred seasonal period for sampling should occur when biological diversity is highest. In general, this will correspond with the seasonal recruitment cycles of benthic organisms that are related to climate and the abundance of food supply. Many insects with freshwater stages reproduce in the spring and fall, although others have multiple cohorts throughout the open water period. For many lotic habitat types, sampling is conducted during the fall (September/ October) when the majority of taxa will be present and/or large enough to be collected by the sampling equipment and flow regimes allow access for sampling. In large lakes where the benthic community is often dominated by annelids, crustaceans, and molluscs, insect emergence periods and hydrologic regimes are of less importance for determining the sampling period.

If historical benthic invertebrate community surveys exist for the system under investigation, it would be useful to examine the data and, if appropriate, conduct the survey during similar periods so that the surveys can be compared. Other factors that may influence the sampling period include seasonal flow disruption such as extreme high or low flow conditions, freezing and ice-scour, mine effluent discharge conditions, the type of sampling gear, and the feasibility of sampling and field crew safety. Sampling during periods when effluent is not being discharged should be avoided. An understanding of the seasonal patterns and life cycles of the taxa along with changes in the hydrologic regime found within the specific system is also helpful to determine the appropriate timing for the survey. Rosenberg and Resh (1993), Johnson et al. (1993), Rees (1984), Malley and Reynolds (1979), Barber and Kevern (1974) and Jonasson (1955) provide information that may assist in the selection of a sampling period.

5.10.2 Sampling Season Considerations for the EEM program phases

For **Initial Monitoring**, it is recommended that efforts be concentrated within a single seasonal sampling period, unless previous data indicate that there is more than one critical time period for the benthic community in the study basin. As this seasonal period will then be used in subsequent studies it is important to make this decision after compiling all available site-specific data regarding taxa life history characteristics and hydrologic discharge regimes.

For **Periodic Monitoring**, reviewing the previous monitoring information and the life history characteristics of the invertebrates collected should confirm if the sampling season used was the most appropriate or sensitive.

Similarly, the sampling season for **Focussed Monitoring** will be the same as previous monitoring unless, upon review of the previous results, there is scientific or logistical justification for a change. Furthermore, additional seasons may be warranted to help determine the magnitude of the response of the benthic community. For example, if the sampling season is during a time that the life stage of a particular invertebrate is not present, then an additional sampling season may be necessary to determine if effects are seen for this specific case. Bivalves are an example of this case, as they are not easily sampled in the fall, which is often a critical period for many other invertebrates. In this case an additional season could be added to the monitoring program with the sampling program designed to answer this site-specific concern (i.e., an additional summer sampling trip where methods designed for bivalve sampling are used).

For most marine or estuarine areas the sampling season could be anytime from spring through mid-fall. For temperate marine environments, benthic sampling is usually conducted in late summer or fall as some benthic forms have planktonic larval stages that do not settle to the benthos until later in the season and populations with spring recruits have stabilized by this time. For Arctic areas, the appropriate time period would likely be late summer or early fall, when the long day-length and warmer temperature have allowed some time for growth and development of flora and fauna and there is no sea-surface ice to contend with. In general, reproductive periods and patterns of abundance of benthic species are related to tidal cycles, season and the abundance of food supply.

5.11 Statistical Considerations for Study Design

This section provides guidance in regards to statistical considerations applicable to the study design of each phase of the EEM program. These statistical considerations recommended include the following sequential steps:

1) selection of the effect size for the power analysis,

2) selection of α and β levels for the power analysis (the same levels are also to be used during the statistical analyses on the compiled monitoring data) and finally,
3) a power analysis to determine the appropriate level of sampling effort at the station scale (i.e., number of stations per area).

In addition, a discussion is included which provides guidance on determining the number of field sub-samples which should be taken at a given station and how this field sub-sample data could be used to improve future study designs.

5.11.1 Setting Effect Sizes for determination of sampling effort

One of the most critical aspects of EEM for mines is the decision on the magnitude of an environmental impact that is deemed to be ecologically important (Mapstone 1995). An estimate of this effect size (ES) must be decided upon *a priori* (i.e., before a monitoring program is designed) in order to determine the sample size to have sufficiently high statistical power to detect an impact of magnitude ES, if it has occurred. Quantitative guidelines providing ES values have not been widely developed for ecological communities, including benthic invertebrate assemblages in either marine or freshwater systems (U.S. Environmental Protection Agency 1990; Elliott 1996, Striplin 1996). In large part, this has resulted from the complex nature of aquatic communities and the difficulty in making quantitative predictions of the expected magnitude of change following an impact, such as might result from the discharge of mining effluents. Despite these difficulties, some attempt is needed at providing initial estimates of scientifically grounded ES values. Note that, although the treatment here of effect size is sufficient in regards to power analyses and the determination of sampling effort, there is further discussion of the different uses of effect size in the Data Interpretation Chapter (Chapter 9).

To determine the sampling effort (i.e., number of stations) there are several approaches to setting effect size outlined below. These approaches are presented in order of ecological relevance although feasibility may be limited by data availability.

Option 1 below would be the preferred (i.e., most ecologically meaningful) choice when extensive background data are available or alternatively, if sufficient, ecoregion-specific data are available, Option 2 is recommended. In the absence of any background information it is recommended that ES initially be set as any change in a measured benthic descriptor that is = 2 SD from reference (Option 3). The use of an effect size other than this will be made with the appropriate scientific argument.

Option 1. ***A priori*** **setting of ES as a fixed percent change based on existing ecological information on the magnitude of an effect known to have long-lasting deleterious effects on the ecosystem.**

Mapstone (1995) gave a hypothetical example of the derivation of ES by using the percent loss in coverage of a key algal species known to have serious, long-lasting deleterious effects on the functioning of the aquatic community being monitored. Unfortunately, ecological information at that level of detail is seldom available (Keough and Mapstone 1995). Extensive previous ecological data are recommended to obtain a full understanding of the ecological meaning of a "detrimental" percent change in the species or environmental descriptor at the site. Therefore, this type of approach would likely not be suitable unless extensive, quality historical data are available. The approach may, however, be possible for some response variables and/or at a later time as ecological understanding of the affected communities improves. At present though, due to the complex nature of most aquatic communities, quantitative guidelines providing fixed ES values have not been widely developed or accepted (see Lowell 1997). One promising approach under development is to incorporate experimental or field data on obviously impaired sites into the decision making process (Robert Bailey, unpublished manuscript, University of Western Ontario, London, ON).

Option 2. ***A priori*** **determination of a standardised, habitat-specific ES calculated from previous data obtained from appropriate reference areas within the same ecoregion.**

This would be developed on the basis of what is considered representative of a "normal" biological community as determined from reference areas across a number of mine sites (or other sites for which comparable reference data are available) within the same ecoregion. As for Option 3 below, ES would be determined from the normal range of variability in the reference areas, using ±2 SD of the reference station values within each reference area. For each response variable, this number would be calculated for each reference area within the ecoregion, and the median of all the calculated area values would be used as a habitat specific ES for that response variable for that ecoregion. The median is used instead of the mean to minimize the influence of outlier (non-typical) reference areas on the calculation. This procedure yields a fixed value for ES (e.g., ±30% taxa), which may or may not be equal to ±2SD for any particular reference area used in the calculation. Care must be taken to include only data from similar habitats for all areas, so as not to artificially inflate the estimate of the ecoregion-wide normal range of variability.

Option 3. ***A priori*** **setting of ES at +/- 2SD calculated from reference area data.**

The suggested approach in the absence of other ecologically rational criteria (see options 1-2 above; and see Lowell 1997 and Environment Canada (1995) is to define ES in terms of the normal range of variability observed in the reference community at the mine site being examined. A commonly accepted criterion for the normal reference condition is ±2 standard deviations (±2 SD) of the reference station values about the reference area mean, assuming the data are normally distributed. Expressed as a percentage of the reference mean, this equals ±2 times the coefficient of variation for the reference mean (±2 CV). This approach makes reasonable ecological sense and appears to have been applied successfully elsewhere (Kilgour 1997). For example, ±2 SD using reference station values has been used to flag departure from normal reference conditions for The Pulp and Paper EEM Technical Guidance Document for Cycle II (Environment Canada, 1998).

5.11.2 Selecting α and β

During any statistical analysis there are four potential outcomes in regards to the accuracy of the conclusions (Table 5.4). For two of the outcomes, the conclusions are correct and, in a properly designed monitoring program, the frequency of these correct conclusions should be far greater than incorrect conclusions. However, there are always potential errors in any statistical evaluations of effects and the two types of possible statistical errors have been defined.

Type I error occurs (at probability α) if the null hypothesis that there is no effect is rejected when in fact it is true (i.e., an area is declared as being different from reference when it is not).

Type II error occurs (at probability β) if the null hypothesis is accepted when it is false (i.e., the area is declared as being the same as reference when it is actually impaired).

The **power** of a statistical test is defined as $1-\beta$, the probability associated with correctly rejecting the null hypothesis when it is false. Therefore, α is the risk to industry, and β is the risk to the environment.

Table 5-4: Potential outcomes from statistical analyses. (✓ = no error, conclusions made are correct)

	Accuracy of Conclusions	
Statistical Conclusion	Ho *is* true	Ho *is not* true
Reject null hypothesis Conclusion $\text{mean}_1 \neq \text{mean}_2$ (e.g., Reference ≠ Exposure)	Type I error (α) *Risk to industry*	✓
Accept null hypothesis Conclusion $\text{mean}_1 = \text{mean}_2$, (e.g., Reference = Exposure)	✓	Type II error (β) *Risk to Environment*

To determine the sampling effort necessary to detect the recommended ES, decisions must be made about the magnitude of Type I and Type II errors that are acceptable. In testing whether an exposed community falls outside the pre-determined guidelines for a "healthy" community, a low probability of a Type I error (α) is usually allowed so that a normal community will not be mistaken for an impacted one. Traditionally, α has been set at 0.05 but there is little literature or consensus on what is an appropriate level for β (Lowell 1997). However, the monitoring program should be designed to provide a reasonably high probability of statistically detecting a predetermined ES if it has occurred. In other words, the power of the test should be high. Recently, because of the potentially high cost (both ecological and monetary) of failing to detect environmental impacts, many benthic impact researchers argue that α should be set at least to the same level as β (Alldredge 1987; Underwood 1993; Mapstone 1995).

We consider it important that α and β be set to the same level and we recommend α and β be set at 0.10 or less (e.g., equal and 0.05 is also acceptable). This recommendation on appropriate α and β levels for Type I and Type II errors have to be made within the context of the power of the test, given that ES has been set using one of the options given above and after reviewing the implications for sampling effort (section 5.11.3). However, in recognition that, the setting of α and β equal at 0.10 is a relatively new concept for Canada's EEM program, the use of other α or β levels is possible with the appropriate scientific justification (e.g., setting more rigorous type II error (β) if the environmental risk is greater). Consultation with the AO is recommended when using ES Options 1 or 2 (above) in cases where power analysis recommends the use of unreasonably high sample sizes.

5.11.3 Determination of Sampling Effort using Power Analysis for ANOVA (and Regression/ANCOVA) designs

The number of stations and their allocation between discreet reference and exposure areas or along a gradient must be determined such that the sampling design will be scientifically defensible and provide adequate data for decision-making for each mine site. The allocation and distribution of these replicates is dependent on the site-specific conditions and the sampling design used. As discussed above, ES and Type I and Type II error must all be taken into account and set *a priori* to determine what sampling effort is necessary using power analysis.

Power analysis is used to determine the sample size necessary to achieve a test adequate to detect an impact with a magnitude = ES, where ES has been determined prior to sampling. Using ES, α, β and some estimate of reference SD, and making some assumptions about the distribution of the data being evaluated, a scientifically defensible sampling strategy can be devised. The discussion below outlines the most basic (i.e., C/I ANOVA) procedure for determining samples size at the station level for testing each specific biotic metric. In cases where the sample size calculated for one variable (e.g., abundance) is greater than that calculated for another (e.g., number of taxa), the greater sample size should be used.

For a C/I ANOVA design, once ES has been determined, the levels of α and β have been selected, and SD for the particular mine site in question has been estimated, they would then be plugged into the power analysis equation to calculate the sample size to detect an impact of magnitude ES between or among areas at a given power level.

It should be noted by the user that, when using ES Options 1 or 2, the determination of the sample size is based on the variability amongst replicates for the reference data. ANOVA analyses assume that the variability amongst replicates for the exposure area is similar to that for the reference area. Although ANOVA's are fairly robust with respect to violation of normality assumptions, if the variance within an exposure area is much higher (or lower) than within the reference area, ANOVA comparisons are not appropriate, unless the variances can be made homogeneous by transformation. For the case where the exposure and reference variances remain significantly different following transformation, the power analysis outlined here may over- or more likely under-estimate the number of sampling stations. Non-parametric tests may be used; however, the estimate of power described herein will no longer be relevant. In some cases, unequal variances may be an artifact of insufficient field sub-sampling (see section 5.11.4).

For a basic C/I ANOVA design, the estimated sample size (number of replicate stations per reference and exposure areas assuming variance is reasonably similar in both areas) to detect a change of ±2 SD at a given level of power can be conveniently calculated by arranging the standard power analysis equation as follows (Green 1989):

$$n = 2(t_\alpha + t_\beta)^2(SD/ES)^2$$

where:

n = sample size
t_α = t value for α significance level (Type I error; two-tailed)
t_β = t value for β significance level (Type II error; one-tailed)
SD = standard deviation
ES = effect size

For ES Option 3, both SD and ES cancel out of the power analysis equation (see Lowell 1997) because ES = 2SD. More specifically, the variability calculated from a preliminary survey cancels out of the equation and, consequently, preliminary data does not provide any additional information needed for calculation of the number of stations to achieve a given level of power.

The power analysis equation can then be modified by expressing ES in terms of d, the number of standard deviations separating the reference and impact means (d = ES/SD), and then substituting to yield (assuming equal reference and exposure variances):

$$n = 2(t_\alpha + t_\beta)^2/d^2$$

So, to determine the sample size to detect an ES of 2 SD, d would be set equal to 2. The equation is solved iteratively by choosing an approximate value of n to look up t_α and t_β and then using the solution to find a more accurate n; the procedure is repeated until arriving at a final estimate for n. An alternative equation, which yields an approximate n in one step, is given as (Guenther 1981; Alldredge 1987; Environment Canada and Department of Fisheries and Oceans 1995):

$$n = (2(Z_\alpha + Z_\beta)^2(SD/ES)^2) + 0.25(Z_\alpha)^2$$

where:

n = sample size
Z_α = standard normal deviate for α significance level (Type I error)
Z_β = standard normal deviate for β significance level (Type II error)
SD = standard deviation
ES = effect size

By using the above power analysis equation, we can examine the effects on sample size of setting α and β at different levels for detecting an effect of ±2 SD (Table 5.5). Pre-calculated tables of n are available for a variety of values of α, β, and SD (Alldredge 1987; Cohen 1988; Environment Canada and Department of Fisheries and Oceans 1995). This implies that, in the recommended case scenario ($\alpha = \beta = 0.1$ and ES = 2SD), the sampling effort will require a sample size of 5 replicate stations, well within the range used in many benthic surveys (Resh and McElvray 1993). Alternatively, if alpha = beta = 0.05 and ES = 2SD, the sampling effort would only increase to 8 replicate stations.

Table 5-5 Sample sizes to detect difference of 2 SD

α	**1-β**			
	0.99	**0.95**	**0.90**	**0.80**
0.01	14	11	10	8
0.05	11	**8**	7	5
0.10	9	7	**5**	4

SD, α, and β as above

Similar considerations and procedures apply to power analysis and determining sample size for regression and ANCOVA monitoring designs (simple, radial, and multiple gradient designs). The reader is referred to the Data Interpretation and Assessment chapter (Chapter 9) and the appropriate literature (e.g., Cohen 1988) for further guidance on these procedures.

5.11.4 Determination of Sampling Effort for RCA designs

The issue of replication is somewhat different when using the RCA. Replication is at the station scale and it has been shown (Rosenberg et al. 1999) that variation within a station is much less than among stations and therefore single samples are taken at stations and variation among stations is used to describe the reference condition. The number of reference replicates is determined by the number of stations in the group to which the exposed station is predicted to belong using the RCA. This is determined when forming the groups of reference stations in the initial classification (see section 5.6) BUT has been set to a minimum of 10 stations. The variation among the reference stations forming the reference group determines the Type I error which has been set at 0.1 by using a 90% probability ellipse. Because, in this approach, single exposure stations are compared to multiple reference stations (minimum of 10), it is not possible to set Type II error which requires an estimate of the variance associated with a single station. A surrogate can be applied by taking more than one sample at the exposed station, but this is estimating within-station error rather than the appropriate variation at the among-station level. Clearly Type II station error cannot be determined when there is only one member of the population of exposed stations. Therefore, the power analyses referred to above would not be applicable for the RCA study design. Consequently, RCA studies should be designed in a way that provides an accurate and precise determination of reference conditions so as to maximize the likelihood of detecting departures from reference conditions at exposure stations, when they exist.

5.11.5 Determination of Sampling Effort for field sub-sampling

The objective of multiple field sub-samples at each replicate station is to ensure that the sampling effort will produce an accurate reflection of all the metrics of interest (e.g., taxa richness, abundance, etc.) for each station that is sampled. This is necessary because species may not be homogeneously distributed throughout a station (which is much bigger than the size of the physical sampling apparatus being used). Inadequate sub-sampling both 1) gives an imprecise estimate of the true mean for each station and 2) can contribute to an inflated estimate of the true among-station variance, thereby decreasing power.

Therefore, the allocation of field sub-samples within a replicate station depends on the following two inter-related factors which should be considered during any benthic sampling design exercise. However, in the absence of background information, the recommended minimum number of field sub-samples to obtain from each station is **three**.

1. The abundance and degree of aggregation of organisms in relation to the desired level of precision for station estimates.

For a given station, the number of field sub-samples needs to be sufficient to give a mean and variance which provides confidence that a representative number of animals has been captured (for review see Burd et al. 1990). The more aggregated a community, the higher the variance of mean abundance for each replicate station. Elliott (1977) and Holme and McIntyre (1984) suggested the same simple method of determining the number of field sub-samples to obtain a predefined level of precision. Elliott (1977) suggests that toleration of an index of precision (D) of 20% (i.e., that the standard error is equal to 20% of the mean) is acceptable for most bottom samples.

The number of field sub-samples can then be calculated as follows:

$$n = s^2 / D^2 \ \bar{X}^2$$

where,

$\bar{X}$ = the sample mean
n = the number of field sub-samples
s^2 = the sample variance
D = the index of precision (i.e., 0.20)

Thus, to determine how many field sub-samples (i.e., grabs) per replicate station will provide an estimate with 20% precision, previous data can be used to determine the mean and variance and hence, the appropriate number of field sub-samples. This determination may vary from location to location along with changes in the mean to variance ratio. It is recommended that the number of field sub-samples be calculated for locations which exhibit the highest variability and that the resultant sample size be applied equally to all areas to standardize sampling effort. Although this recommendation will produce better precision in the less variable habitats, it is a conservative approach and maintains equal sampling effort between areas and replicate stations. Also of note is that, with aggregated populations, although the overall mean should remain the same, depending on the scale of the aggregation in relation to sampler size, variance will change with the size of the sampler. Therefore, sample size estimates using preliminary data are only relevant to a sampling program which would employ the same type and size of sampler with which the preliminary data were obtained. In cases where this sampling effort cannot be determined from a previous cycle's data, counting organisms in field sub-samples from the current survey as they are processed and calculating means and variances will allow determination of how many grabs should be processed in the laboratory. However, this *a posteriori* approach necessitates that a sufficient number of grabs was obtained during the field survey in the first place, so the effort to calculate the sample size within a replicate station *a priori* should minimize problems with insufficient sampling effort.

A related approach uses abundance and variance to determine sub-sampling effort and precision and can be used for determining the number of field sub-samples at all replicate stations. It is derived from the relationship between within-station mean abundance and variance across all replicate stations in the area or gradient being evaluated. Downing (1979, 1986) used Taylor's power law (1961) to estimate aggregation in a freshwater benthic community, and thus sampling effort required to reduce variance to an acceptable level. In effect, a given number of organisms is required for each replicate station in order to produce the precision of within-station mean abundance from sampling. Vezina (1988) used the same approach to determine empirically the degree of aggregation inherent in marine benthic communities. This approach entails calculating a power regression equation that describes the log/log "relationship" between within-station mean abundance and variance across all stations; this provides a formula which is then used to determine the estimated variance expected for a given abundance of organisms in that survey region. From this, the estimated variance for each mean abundance at each replicate station is calculated and then used in the same way as Elliott (1977) to estimate the number of field sub-samples at that replicate station. The difference between the two methods of Elliott (1977) and Downing (1979) is that the variance used in the equation to determine the number of field sub-samples is based respectively on 1) sample variance and 2) variance calculated for the sample mean from the power regression equation for all the samples in the survey region. Furthermore, the index of aggregation (slope) from the power regression equation can then be used to determine the most appropriate data transformation for parametric statistical analyses. Unfortunately, this method assumes that the overall assemblage has a uniform aggregation throughout the study area, which may or may not be true when an external environmental stressor is applied. However, the degree of goodness of fit of the mean and variance data to the log regression equation provides a good indication of how true the homogeneous aggregation assumption is. If there are extreme outliers, they should be taken out of the analysis to avoid skewing the results. Because the aggregation of benthic communities can change as environmental conditions change either naturally or unnaturally, it is wise to review the relationship between mean and variance every time benthic samples are collected. Finally, it should be noted in the above discussion that the "power regression equation" used here to calculate the number of field sub-samples is unrelated to the "power analyses" used to determine the number of replicate stations discussed in the previous sections.

2. The number and distribution of different species in relation to obtaining a representative collection.

Determination that sufficient species have been sampled can be done using simple rarefaction methods such as the "species abundance curve" or species/ sampling area curve (for review see Burd et al. 1990) which compare the number of species obtained versus number of individuals for different numbers of pooled replicates. This analysis is particularly important in arctic areas, where diversity may be high, but only on a geographic scale much larger than is feasible to sample (i.e., number of species relative to abundance is high, but abundance is quite low – this can also occur in the deep-sea). Because of assumptions inherent in the underlying distribution of fauna related to logarithmic species abundance curves, a more sophisticated approach is the "similarity/ sampling area" curve, which uses similarity indices on presence/absence data to determine the sampling effort to obtain an acceptable overall faunal similarity between replicate stations (Weinberg 1978, Kronberg 1987).

If preliminary data are unavailable or unsuitable for determination of the number of field sub-samples to obtain a representative collection of species, a check on sampling effort could be performed very simply. If it is estimated that X number of grabs per replicate station are sufficient to achieve a data quality objective of retrieving 95% of benthic species present at any replicate station, more grabs can be collected at a few selected replicate stations and analyzed. Determination of a taxa richness plateau from these extra samples determines whether the number of grabs were sufficient to achieve the 95% objective (using a species area curve).

5.11.6 The Use of ordination probability ellipses for RCA designs

A large water quality survey on rivers conducted in the UK in 1990 provided the impetus for the development of methods to circumscribe the continuum of responses into a series of bands that represented grades of biological quality (Clarke et al 1992). The study produced a simplification of the continuum of responses in sites ranging from good to poor biological quality. It was seen as an appropriate mechanism for obtaining a simple statement of biological quality which allows broad comparisons in either space or time that are useful for management purposes. From a management perspective it is desirable to assign a degree of impairment. This can be done by setting response categories from mild to severe impairment. In the study by Clarke et al. (1992), a number of schemes for categorizing the response were considered and tested. The threshold between unstressed and stressed sites (band A) was set at the 90% probability level (SD = 1.64) for number of taxa and the BMWP score and 95% for the average score per taxon (ASPT). In Australia the threshold is set at 2 SD from the reference site mean for the number of taxa. Finally, 95% is frequently set as the limit for determining a biological effect for univariate data and single community descriptors (Lowell, 1997). The strategy employed in the UK (Wright 1995) to discriminate between degrees of impairment was to quantify the thresholds for stressed and non-stressed sites via the setting of three equal size bands, as Wright (1995) argued that there was no logical basis for an alternative scheme for dividing up the continuum of sites.

A similar approach can be adopted for defining degrees of impact using multivariate ordination. The reference invertebrate assemblage can be described by its distribution in ordination space, and the assemblage at any given site is characterized by its position in that XY space (Figure 5.9). The greater the similarity between sites the closer together they are in XY space. Using this approach to set effect size for an invertebrate assemblage, all the reference sites are plotted in XY space together with a test site(s). The likelihood of the test site being the same as the reference sites is quantified by constructing probability ellipses for the reference sites ONLY. Reynoldson et al. (1999) selected the 90% probability ellipse as representing the first band, the threshold for a site being considered equivalent to reference. The rationale for using the 90% ellipse rather than the more typical 95% was based on the fact that a multivariate approach will tend to be noisier than individual univariate measures and therefore a more conservative threshold was deemed appropriate. Sites located in ordination space inside this smallest ellipse (90% probability) would be considered as equivalent to reference and therefore unstressed. Two other probability ellipses are used (Figure 5.9), that are equal in width, to describe further divergence from the reference state, following the argument used by Wright and co-workers (Clarke et al. 1992, Wright 1995). Sites between the smallest (90%) and next ellipse (99% probability) would be considered possibly different; there is a 1 in 10 chance that sites will fall in this band through normal variability. Sites between the 99% and the largest ellipse (99.9% probability) are considered different, there is a 1 in 100 chance that these sites would incorrectly be described as different. And finally, sites located outside the 99.9% ellipse are designated as very different.

5.12 Benthic Invertebrate Taxonomy

5.12.1 Taxonomic Level of Identification

Identification of organisms in a community must be adequate to meet the objectives of the assessment program. Recent research indicates that family level identification provides sufficient taxonomic resolution to detect community responses to human disturbances (Warwick 1988a, 1988b, Bowman and Bailey 1997). As discussed below, the level of taxonomic resolution used may necessarily vary across the different monitoring phases with finer taxonomic resolution needed to detect more subtle environmental impacts.

The recommended level of taxonomic identification is ***family*** for **Initial and Periodic Monitoring for freshwater**. All summary statistics and descriptive metrics should be calculated and reported at the family level for submission for the Initial and Periodic Monitoring interpretative reports. Organisms that cannot be identified to the desired level of taxonomic precision should be reported as a separate category in the fundamental data set. It is recommended that investigators use taxonomic keys appropriate to the geographic region of study; a detailed list of taxonomic references for various groups of freshwater organisms is provided in section 5.22.

Although, for data reporting and determination of effects the recommended level of identification is at the family level, mines can, if desired and cost effective, proceed with benthic invertebrate identifications to a lower level. There may be site-specific conditions that warrant a lower taxonomic level for some or all familial groups. For example, historical benthic invertebrate information may be identified to lower taxonomic levels, and it may be desirable to identify subsequent surveys to a similar level for comparative purposes. If a lower taxonomic level has been used, either in historical data or during a current survey, the summary statistics and descriptive metrics can be reported at this level, provided a summarized data set at the family level is also included.

The two objectives of the **Focussed Monitoring** survey may require different levels of taxonomic resolution. Determination of the *geographical extent of the effect* may be addressed adequately with ***family*** identification. This first objective is similar in scope to Initial and Periodic Monitoring, the major difference being the addition of exposure areas further from the effluent discharge. Family level identification would provide the information necessary for calculating and reporting the required summary statistics and descriptive related to extent of the effect.

The second objective of Focussed Monitoring, *determining the magnitude of the effect,* may use family identification, or, it could warrant investigation at a lower taxonomic level. The question of magnitude of the effect with regard to taxonomic level can be addressed by the following question:

- What is the magnitude of the effect on specific taxonomic groups that may be sensitive to the site-specific mine effluent characteristics (e.g., how many groups within a sensitive family are affected)?

Addressing the magnitude of effect during Focussed Monitoring can be accomplished by using one of the options outlined below:

- Identify all samples collected during Focussed Monitoring to the Lowest Practical Level (see section 5.22). Establishing the magnitude of effect in this way provides additional information that may be useful for the study design exercise at the onset of Investigation of Cause.
- Re-analyze families that were significantly affected during Initial Monitoring to identify indicator taxa that can be used to assess magnitude of effect at further afield stations. For example, if an effect during Initial Monitoring was observed for the Family Baetidae (Order Ephemeroptera), all Baetidae could be identified to a lower level (e.g., Genus) for the Focussed Monitoring program. This approach would catalog the "sensitive" taxa within the family, and the magnitude of effects would be established by examining this subset of sensitive taxa.
- Other scientifically defensible approaches may be used to identify magnitude of effect as required for Focussed Monitoring.

5.12.2 Reference Collections

Consistency in taxonomic identifications within and between surveys is essential to obtain useful information environmental effects monitoring. Therefore, for comparative purposes and quality control of taxonomic identification, EEM recommends the maintenance of reference collections of organisms. In addition, it is recommended that when the reference collection is established, a professional taxonomist verify the identification of each taxon. Reference collections have several benefits including their use for confirming identifications, ensuring consistent taxonomy between surveys, and in the training of personnel. Protocols for establishing and maintaining reference collections for benthic invertebrates are detailed in a report prepared for DOE's Fraser River Action Plan (Green 1994). Following the recommendations of this report, each mine (or group of mines) should compile and archive a complete reference collection with several specimens of representative-sized individuals for each taxon. The collection should encompass representative organisms from each area in the survey, be labeled according to the location and date of collection and updated as appropriate (i.e., when a taxon is collected). This type of reference collection will not occupy a large space, a rather small cupboard should be sufficient, and should be in the custodianship of the mine. If a mine does not have the facilities or personnel to maintain their own reference collection, universities or museums may be willing to fulfill this function. Whether the collection is housed on or off site, clearly it is in the mine's interest to maintain control over the reference collection for access and comparative purposes during future surveys.

In **marine/estuarine** environments, it is recommended that all benthic invertebrate organisms be identified to the ***lowest practicable taxonomic level (LPL).*** For submission of interpretative reports, all summary statistics are calculated and reported to the LPL. For marine and estuary habitat, currently, the taxonomic evaluation of benthic samples should be carried out to the lowest taxonomic level readily possible - genus and species wherever feasible. Various authors have examined the utility of using higher taxonomic classifications for environmental monitoring of organic pollution sites in Europe (c.f. Warwick and Clarke 1993 and references therein). However, the consideration of using higher taxa levels in marine benthic studies requires more study and specific attention to the effects of metal mines on benthic biodiversity. For marine benthos, juvenile or non-adult fauna should be identified and enumerated separately from adults, as they show different patterns of response to environmental effects

For marine samples it is suggested that, if sufficient numbers of specimens are available in the reference collections, they could be used for a further purpose; to develop a size and biomass database for each mine (i.e., as another indicator or tool, see section 5.24). For these purposes, 5-10 representative specimens per species are recommended with mean width, lengths and blotted wet weights recorded for each group.

5.13 Supporting Environmental Variables

Supporting environmental variables are useful to provide explanatory evidence for data interpretation, and the visual habitat mapping exercise recommended in Site Characterization should be expanded during the **Initial**, **Periodic** and **Focussed** Monitoring surveys. In addition to the study area characteristics documented during Site Characterization (Chapter 3, Table 3.1), appropriate environmental supporting variables should be measured at the time of the benthic survey. If scientifically warranted and logistically feasible, a limited set of variables can be measured more frequently (e.g., water temperature).

Several useful sources of information on selecting habitat descriptors to explaining patterns in lotic invertebrate communities include Wright et al. (1984), Corkum and Currie (1987), Corkum (1991) and Ormerod and Edwards (1987). For lentic communities, articles by Johnson et al. (1991), Reynoldson et al. (1995) and Bailey et al. (1996) provide lists of the most useful habitat descriptors. Additionally, several authors have described responses in invertebrate communities and demonstrated the ability to distinguish natural variation from anthropogenic stressors (Clarke 1993; Clarke and Ainsworth 1993; Hargrave et al. 1993, 1997; Warwick 1993; and Warwick and Clarke 1993).

Supporting environmental variables can be divided into two categories:

1) Exposure variables that can provide an indication of the extent of effluent exposure, and
2) Explanatory variables, which can assist in the interpretation of the monitoring results, and allow investigators to determine if natural environmental variables influence, or are related to, the observed benthic invertebrate community structure.

Exposure variables are those variables measured as part of water quality monitoring. Chapter 6 lists other water quality parameters that may be useful to measure on a site specific basis. The water quality monitoring variables will be measured within each sampling area at the time of the benthic invertebrate community survey.

Explanatory variables should include appropriate, site-specific variables from each of the broad categories outlined in Table 5.5. These should be measured quantitatively at each replicate station during the benthic survey. Explanatory variables should be measured at each station in order to allow the linking of environmental correlates with invertebrate community structure. Although this level of sampling effort for explanatory variables is recommended for all EEM phases and study design options (section 5.6), it is particularly useful for gradient and reference condition approaches.

Table 5-6: Recommended supporting explanatory variables for freshwater benthic invertebrate habitats.

These supporting variables should be measured at each station during all EEM phases. Methods are described in section 5.21.

Category	Variables	Notes
Habitat Description	Stream order Drainage area	
Physical-Chemical Measurements	Water temperature Dissolved oxygen pH Conductivity Alkalinity Turbidity Nutrients	(Installation of a temperature logger to allow measurement of total degree-days would be more useful than spot temperature readings.)
Morphometric Measurements	Bankfull width Wetted width Depth Slope	See section 5.21 for definitions of widths
Flow Characteristics	Velocity Stream hydrograph	Seasonal hydrographs would be instructive for selecting appropriate seasons but it is assumed that their use will be limited to those sites where the information is available from a local EC water survey station. The installation of a discharge gauging station is not recommended nor are single point discharge measurements unless there are scientific reasons to do so.
Substratum Characteristics	Embeddedness Substratum composition (% cobble, gravel, sand, etc.)	For erosional habitats For erosional habitats
Sediment Chemistry	Fine sediment particle size and TOC	For depositional habitats *Note – samples can be archived for later chemical analysis if an effect is observed*
Riparian Zone Characteristics	Riparian vegetation Canopy cover (%)	

5.14 Integrating Benthic Invertebrate and Sediment Programs

Sediment samples must be collected at the same replicate stations and at the same time as benthic invertebrate sampling. Samples are analyzed for sediment particle size distribution and total organic carbon with representative subsamples archived for possible, future chemical analysis. See Chapter 8 for guidance on recommended parameters and methods of analysis.

For monitoring programs where the sampling of benthic invertebrates is conducted in erosional habitat, sediment sampling is not recommended as a standard supporting environmental variable. Some methods for retrieving sediments from erosional zones require elaborate equipment or two field visits, one for the placement and one for the collection of sediment traps. However, site specific conditions may warrant the consideration of sediment sampling in some erosional habitats as useful information regarding exposure can be obtained with these methods (see Chapter 8). These approaches could be considered during the study design exercises for Focussed Monitoring or Investigation of Cause as an additional supporting variable or tool for determining cause.

5.14.1 Review of the Supporting Environmental Variables for Subsequent Monitoring

A review of the previous monitoring results and environmental variables measured should be performed during the study design for **any subsequent monitoring** with consideration of the objectives of the specific phase. Assessment of previously measured environmental variables should include addressing the following questions:

- Were the measured supporting environmental variables correlated with the benthic invertebrate community characteristics?
- Are there any supporting environmental variables that should be added to the list? Are there some variables that are unnecessary to measure?
- Was the degree of exposure to mine effluent determined or verified in the study area?

Sources of information on selecting habitat descriptors for explaining patterns in marine invertebrate communities are Clarke 1993; Clarke and Ainsworth 1993; Hargrave et al. 1993, 1997; Warwick 1993; and Warwick and Clarke 1993. These authors have described responses in invertebrate communities and demonstrated the ability to distinguish natural variation from anthropogenic stressors. Not all factors will be applicable to all habitats types. For example, salinity may only be important in intertidal or estuarine habitats. Sediment particle sizes, sulfides and TOC are only applicable in unconsolidated sediments.

Table 5-7: Recommended supporting explanatory variables for marine/estuarine benthic invertebrate habitats.

These supporting variables should be measured at each station during all EEM

Category	Habitat$^{(1)}$	Variables recommended at all sites	Site specific variables $^{(2)}$	Notes
Habitat Description	E,M		Water column stability	Stagnation, bottom resuspension, turnover due to surface ice, etc.
	E, I	Wave exposure	Ice scour, freezing Tidal cycle and range	
Physical-Chemical Measurements	E, M	Dissolved oxygen near bottom	Water temperature Color or Turbidity	Naturel low oxygen can affect communities Strong temperature gradient estuaries can affect communities Beware of seasonnal temperature effect in arctic mine areas Light reduction due to effluent colored or fine suspended particulates may reduce primary production and/or affect filtering apparatus of benthic organisms
	E, I	Light		
	E, M, I	Salinity		
Morphometric Measurements	E, M	Depth		
	E, I	Slope		
Sediment Chemistry	E, M	Sediment Total organic carbon (TOC) or Sediment Volatile Residue (SRV)	Sediment C/N ratio	Effects may be related to decreases in natural TOC from inorganic tailings C/N ratio is affected by terrrestrial (wood) versus aquatic plant sources of marine deposition and useful to eliminate confounding factors such as log booms, pulp mills, etc. Sediment redox provides an indication of the anoxic/oxic boundary in sediments Sulfides in sediments indicate the extent and nature of microbial response to organic enrrichment
	E, M, I	Sediment Eh (redox) Sediment Total Sulfides		
Substratum Characteristics	E, M	Sediment Particle size Bed structure Matrix: framework ratio, degree of embeddedness, texture, colour and thickness layers in cores		
Flow Characteristics	E, M		Current velocity	

(1) E=estuarine, M=marine,I=intertidal

(2) variables recommended if site-specific conditions dictate

5.15 Other Supporting Environmental Variables

Biotic indicators other than benthic invertebrates are not required for assessment during the EEM program. Nevertheless, consideration of other indicators for this program phase are possible if there are site specific conditions which warrant it.

After review of the Initial and Periodic Monitoring results, and during the study design phase of **Focussed Monitoring or Investigation of Cause**, consideration of additional biotic indicators can be determined by assessing the ecological relevance of the following biotic indicators.

- Periphyton or Phytoplankton biomass and taxonomic composition;
- Macrophyte biomass and taxonomic composition;
- Zooplankton biomass and taxonomic composition (for lakes);
- Other measures of benthic invertebrate communities such as, biomass, lower level identification, secondary production and population fitness parameters; and,
- Other scientifically defensible indicators.

The applicability of these biotic indicators to the EEM program is discussed in section 5.24. However, it should be noted that these biotic indicators are potential tools and not alternatives to benthic invertebrate community surveys. Additional tools or indicators may be added as the technology is developed.

5.16 Data Assessment and Interpretation

5.16.1 Data Handling Methods

5.16.1.1 QA/QC for Data Input and Verification

After data entry, the first step in data analysis is to check for transcription errors. Failure to do this invalidates further analyses. All computer entries must be verified by checking a hard copy of the file against the raw data sheets. Someone other than the person who originally entered the data should do this cross-checking. Ecological community studies usually generate large data matrices when an observation on each taxon is recorded at every sampling station for hundreds of taxa. Computers and multivariate statistical programs generally assist in synthesis of the data to elucidate geographic and temporal patterns of community health (e.g. Hensler 1984; Ludwig and Reynolds 1988). Due to the nature and quantity of data, particular care is needed in data handling, to guard against data entry errors or subsequent misuse of data. Double entry systems and transcription checks against the original data records are useful QC techniques. Missing data should be clearly distinguished from a taxon absence by use of unique non-zero missing value codes with code definitions built into each file. Read-only files help to ensure data integrity. QA/QC concerns regarding data analysis include data verification and validity, repeatability and robustness of statistical analyses, and rigor and defensibility of analysis. Gibbons et al. (1993) suggest that another investigator should be able to arrive at the same conclusions if they were to use the methods and data set found in the the report. Other considerations regarding the data verification and analyses are listed below:

1) use trained and experienced personnel;
2) conduct screening exercises to identify transcription errors, outliers and other suspicious data points;
3) provide raw data in a electronic database format and appendices to reports which summarize the data;
4) document the methods (specific statistical tests) and software (if used) used for analysis; and
5) maintain detailed notes regarding the analyses of the survey data.

For further information regarding all aspects of QA/QC procedures for benthic invertebrate programs refer to the 1999 AETE report (Beak 1999).

5.16.1.2 Dealing with outliers

Assuming the data are entered correctly, data should be summarized, screened for erroneous values and outliers, assessed for normality, and transformed if necessary (Gibbons et al. 1993). Visual screening techniques such as Box and Whisker Plots, Normal Probability Plots and Stem-Leaf Diagrams can be used to identify extreme values (true outliers and/or data entry errors) (see Tukey 1977). Norris and Georges (1993) recommend examining abundance estimates for each taxon to determine if numbers are reasonable. They also recommend calculating means and standard deviations because aberrantly high or low values can indicate errors. Extreme values or outliers, which are not errors of some kind, should not be removed from the data set because this will result in the loss of an observation and a loss of power to the benthic invertebrate community survey. Instead, extreme values should be identified in the report and the influence of the extreme value(s) on the results should be determined by re-analyzing the data without the extreme value.

5.16.2 Unknown and Immature Organisms

Some samples may contain immature individuals that cannot be identified to the recommended level of taxonomic precision. A similar situation could also occur when samples are improperly preserved and identifying features are destroyed (e.g. mollusc shells dissolve due to unbuffered formalin). For the purposes of correctly reporting the raw data, these unidentified taxa and their abundances should be provided within the electronic raw data and report appendices. However, for data analysis, investigators must decide whether or not to apportion the unknown individuals according to the ratio of known specimens. This assumes that the ratio of unidentified specimens is similar to the ratio of identified specimens, which may or may not be true. The choices include:

1. not incorporating immature or damaged forms at all,
2. pooling all specimens (i.e., mature/immature, identified/unidentified) and lumping them into one category at the next higher taxonomic level or
3. keeping unidentified taxa as a separate category in the analysis.

Option (1) is not preferred if the "problem" taxa represent a large proportion of the total benthic invertebrate community. Option (2) assumes that all taxa within a higher taxonomic level respond the same way to effluent-related stressors, which may or may not be true. Option (3) will have variable effects on data interpretation depending on the abundance of unidentified taxa. Whatever choice is made will depend on the expertise and experience of the individual investigator; however, it should be fully documented in the methods section of the interpretative report.

For marine surveys, it is recommended that immature and juveniles be counted and enumerated separately from adults whether or not they can be identified to species, so that the adult assemblage can be analyzed without the confounding influence from transient juveniles. Thus, data analyses should show results with and without immatures included. This is because newly settled benthic forms have different survival characteristics than adults, which have been present in the sediments much longer and integrate the effects of habitat perturbations over time. Depending on the timing of sampling, newly settled juveniles may be abundant in samples, but may all die within days due to habitat stressors, predation or competition. This is not to say that data on immatures are not important. Dramatic variations in immature settlement between nearby samples within physically homogeneous habitats may be indicative of varying levels of stressors. It is just important to avoid confounding of results by mixing the groups together for analysis.

5.16.3 Data Reductions and Transformations

Data transformations are often used without consideration to the effects they have on the interpretation of results. Transformations should be applied only with a complete understanding of their effect on the data and its interpretation and only if they are necessary to aid in statistical analyses. Transformations should:

1. make heterogeneous variances homogeneous or make the variance independent of the mean abundance for parametric analyses;
2. normalize distributions;
3. linearize relationships among variables;
4. reduce the effects of extremely dominant taxa within a data set on a multivariate summary (or ordination); and,
5. reduce the analytical problem of too many zeros in a data matrix (see Clarke and Green 1988).

Data reductions should be done only to aid in statistical or multivariate analyses, and for the same reasons as data transformation (see below). Data reductions can include eliminating or rolling up rare taxa, or reducing field sub-samples by pooling or averaging. Protocols for data reductions for marine communities are varied, but subsequent interpretations of data analyses should take these reductions into account. For example, elimination of rare taxa may result in the elimination of 90% or more of the biomass within a given station if those rare taxa are large. In some cases, rare taxa are rolled up into higher groups, which prevents loss of information but adds assumptions about the uniform behavior of mixed taxonomic groups. Reviews of standard methods of data reductions are given in Stephenson and Cook (1980) with some ecological consideration in Burd et al. (1990).

Logarithmic transformations have often been used for benthic invertebrate data because organism abundance typically varies exponentially (Green 1979). A log transformation will reduce the importance of the numerically dominant members, and will improve the likelihood of resolving structure when differences are due to medium-abundance or rare taxa. However, a log transformation is quite extreme. Other researchers have advocated the use of other geometric conversions such as square root, cube root, fourth-root, natural log, etc. (for reviews see Hoyle 1973, Tukey 1977 and Hoaglin et al. 1983, Downing 1981). Downing (1979) showed empirically that the best overall transformation for stabilizing variance in freshwater benthos was the fourth-root ($x^{0.25}$), because this greatly improves the performance of parametric multivariate methods such as ordinations. Vezina (1988) repeated the exercise for marine subtidal communities, concluding that they were empirically less aggregated than freshwater, and require a less extreme transformation (e.g., $x^{0.4}$). However, both researchers emphasize that the mean and variance relationship of any given community must be analyzed to determine the most appropriate transformation. In this way it is possible to check whether or not the transformation used has stabilized the variance.

5.17 Data Reporting Guidelines

Data are submitted in the electronic database format and in hard copy (the interpretative report), as outlined and provided by Envirnoment Canada. The complete fundamental data set, including rare and highly variable taxa and ambiguous identifications should be stored in this manner, even if data filtering has been applied prior to calculation of community descriptors. In this way, other approaches to data filtering, calculation of community descriptors, and analysis can be employed in re-analysis or meta-analysis. A list of relevant details for the field, laboratory and data analysis components of EEM benthic invertebrate survey that should be included and submitted with the interpretative report is provided below.

Field reporting

1. field sheets should be retained for six years;
2. replicate station location (grid coordinates);
3. date and time of sampling;
4. field crew members;
5. habitat descriptions including measures of the supporting environmental variables; and,
6. sampling method used including type and size of sampler and sieve or mesh size.

Laboratory reporting

1. bench sheets should be retained for six years;
2. raw data reported for each individual or pooled field sub-sample, listing taxa present and numbers of individuals;
3. method and level of sub-sampling applied in the laboratory sorting process;
4. sorting efficiency achieved;
5. taxonomic authorities used; and,
6. location of reference collection and report on taxonomic verification.

Data analysis reporting

1. tabular listing of the taxon abundance in each sample as an appendix;
2. tabular summaries of calculated descriptors with variance estimates;
3. estimates of power obtained for the survey;
4. effects of outliers or extreme values on the results (if any); and,
5. a summary of adherence to data quality objectives, standard operating procedures and sampling protocols, and identification of any QA/QC problems.

5.18 Summarizing Monitoring Results

Summarizing data for each monitoring survey should consist of:

A description of the sampling program, including:

- a summary of the study design (and any deviations from the study design)
- geographical description of the station(s) (latitude, longitude and written descriptions)

- number of stations
- number of field sub-samples taken
- type and size of sampler(s) used
- mesh size(s) of sampler and sample reduction or laboratory sorting sieves
- date of sampling trip(s)
- habitat type(s)
- environmental variables measured
- laboratory sorting procedures (e.g., was sub-sampling used)
- QA/QC procedures

A summary of the following Benthic Invertebrate Community Descriptors

- Total invertebrate density*
- Taxon (i.e., Family) richness*
- Simpson's diversity index*
- Bray Curtis index*
- Evenness
- Taxon (i.e., Family) density
- Taxon (i.e., Family) proportion
- Taxa (i.e., Family) presence/absence

* calculations used to determine effects on the benthic invertebrate community.

Many other benthic invertebrate descriptive metrics are available in the literature and are used to address a wide range of questions regarding the benthic invertebrate communities. If desired, additional site specific descriptors may be calculated and used for evaluation of effects. For guidance on selecting these optional descriptive metrics and discussion of their applicability please refer to the reviews by Resh et al. (1995).

5.18.1 Calculation and Reporting of Benthic Invertebrate Community Descriptors

Total invertebrate density: The total number of individuals of all taxonomic categories collected at the station expressed per unit area (e.g., numbers/m^2). Values should be reported for each station, as well as the arithmetic mean (± standard error (SE), ± standard deviation (SD)), median, minimum and maximum for the area.

Taxon (i.e., family) richness: The total number of different taxonomic categories collected at the Station, and the arithmetic mean (±SE, ±SD), median, minimum and maximum for the area.

Simpson's diversity index: Simpson's diversity index (D) takes into account both the abundance patterns and taxonomic richness of the community. This is calculated by determining for each taxonomic group at a station, the proportion of individuals that it contributes to the total in the station. D for each station and arithmetic mean(±SE, ±SD), median, minimum and maximum D for the area must be reported. Simpson's diversity index is calculated as (Krebs 1985):

$$D = 1 - \sum_{i=1}^{s} (p_i)^2$$

where:

D = Simpson's index of diversity;
S = the total number of taxa at the station;
p_i = the proportion of the i^{th} taxon at the station;

Evenness (Equitability): Evenness (E) can be quantified for each station and mean E (±SE) for the area should be reported. Evenness is calculated as in Smith and Wilson (1996):

$$E = 1 / \sum_{i=1}^{s} (p_i)^2 / S$$

where:
E = Evenness
p_i = the proportion of the i^{th} taxon at the station
S = the total number of taxa at the station

Taxon (i.e., family) density: The number of individuals of each family expressed per unit area (e.g., numbers/m^2). Values should be reported for each taxon at each station, and as the mean (±SE) of each taxon for the area.
Taxon (i.e., family) proportion: The percentage abundance for each taxon at each station, and the mean (±SE) percentage abundance of each taxon for the area.

Taxa (i.e., family) presence/absence: A matrix indicating the presence and absence of each taxa at the sampling stations should be reported. The matrix will consist of stations (columns) and taxa (rows).

Bray-Curtis (B-C) Index :

The B-C Index is a distance co-efficient that reaches a maximum value of 1 for two sites that are entirely different and a minimum value of 0 for two sites that possess identical descriptors. These distance coefficients measure the amount of association between sites and the B-C Index is a member of the class of distance coefficient known as a semimetric that some prefer to call dissimilarity coefficients. The B-C Index measures the percentage of difference between sites (Legendre and Legendre 1983) where the distance statistic, is calculated as below:

$$B-C = \frac{\sum_{i=1}^{n} |y_{i1} - y_{i2}|}{\sum_{i=1}^{n} (y_{i1} + y_{i2})}$$

and:
B-C = Bray-Curtis distance between sites 1 and 2;
Y_{i1} = count for species i at site 1;
Y_{i2} = count for species i at site 2;
n = total number of species present at the two sites.

The Bray-Curtis distance (B-C) from a calculated reference median should be reported for each station, and the arithmetic mean (±SE, ±SD), minimum and maximum B-C distance is reported for the area. As the use of this index for determination of effects may be novel to some, a brief literature summary and a detailed example is provided below regarding this specific use of the B-C Index in the benthic invertebrate community survey.

Most of the invertebrate community statistics discussed above are measures of total abundance and taxon richness and do not provide any quantitative information on what kind of organisms are present. A *similarity index* is also recommended as these summarise the overall difference in community structure between reference and exposed sites in a single number, they require no preconceived assumptions about the nature of the community and they only vary in one direction (Taylor and Bailey 1997). Of the various indices available many reviewers have indicated that the Bray-Curtis Index (Bray and Curtis 1957) is the most reliable (Pontasch et al. 1989, Marchant et al. 1984, Jackson 1993, Bloom 1981). Faith et al. 1991 showed that for detecting effects of uranium and gold mines the Bray-Curtis (B-C) Index was superior to seven other indices and provided consistently high statistical power because of its low susceptibility to temporal variability at reference sites and sensitive response to disturbance. It is also unaffected by the nature of communities being compared (Bloom 1981) and differences contribute the same to the Bray-Curtis (B-C) index regardless of the species being rare or abundant. Bloom (1981) showed that of four indices examined only the B-C index reflected accurately the true resemblance over its range.

Example of Bray-Curtis Index for use in the EEM program

To illustrate how this index should be used for the evaluation of effects for the benthic invertebrate component of the EEM program, the following steps with an example data set is presented where 5 stations were sampled from an exposure and a reference area with a total of 5 taxa found to be present.

1) The taxon density is input into a table and,

2) for the reference stations the median taxon density is determined (see example below).

	Taxon density				
Reference Stations	Taxa 1	Taxa 2	Taxa 3	Taxa 4	Taxa 5
Ref 1	2	3	2	3	1
Ref 2	3	5	2	4	3
Ref 3	9	1	1	1	1
Ref 4	4	6	3	4	1
Ref 5	5	4	2	3	2
Reference median	4	4	2	3	1

3) A similar table is constructed for the exposure stations without the median calculation.

	Taxon density				
Exposure Stations	**Taxa 1**	**Taxa 2**	**Taxa 3**	**Taxa 4**	**Taxa 5**
Exp 1	23	4	2	10	1
Exp 2	12	2	2	8	3
Exp 3	14	6	1	6	2
Exp 4	13	1	3	12	2
Exp 5	15	3	2	4	1

4) The distances of each station (reference and exposure) from the reference median is calculated as the following example for reference station 1 illustrates. For this approach, the reference median for a particular taxa becomes y_{i2}, the taxon count for site 2 in above equation.

	Taxa 1	**Taxa 2**	**Taxa 3**	**Taxa 4**	**Taxa 5**
Ref 1 (y_{i1})	2	3	2	3	1
Reference median (y_{i2})	4	4	2	3	1

$\mid y_{i1}-y_{i2}\mid$ or Ref 1- reference median	2	1	0	0	0

($y_{i1}+y_{i2}$)	6	7	4	6	2

Substituting into the B-C equation gives:

$$B\text{-}C_{ref1-refmedian} = \frac{2+1+0+0+0}{6+7+4+6+2} = \frac{3}{25} = 0.12$$

5) The B-C distance from the reference median is calculated for each station in this manner and, should be reported for each station, along with the mean (±SE) for the Area. The sample data set would result in the following B-C distances:

Station	$\Sigma \mid y_{i1}-y_{i2} \mid$	$\Sigma (y_{i1}+y_{i2})$	**B-C distance from median**	**Mean ± SE**
Ref 1	3	25	0.12	0.18 ±0.06
Ref 2	3	31	0.16	
Ref 3	1	27	0.41	
Ref 4	4	32	0.13	
Ref 5	2	30	0.07	
Exp 1	26	54	0.48	0.43 ± 0.03
Exp 2	13	41	0.41	
Exp 3	15	43	0.40	
Exp 4	17	45	0.51	
Exp 5	11	39	0.33	

6) Finally, for the purposes of determining an effect at the exposure area, the mean B-C distance between the reference stations and the reference median (0.18 ±0.06) can be compared statistically to the mean distance between the exposure stations and the reference median (0.43 ± 0.03).

5.19 Evaluation of Results

5.19.1 Is there an effect on the benthic invertebrate community?

The objective of the benthic invertebrate component of the EEM program is to answer the following question:

"Is there an effect on the benthic invertebrate community?" where an effect on the benthic invertebrate community is defined as a statistical difference between benthic invertebrate community measurements taken in an exposure area and reference area of a statistical difference between sampling areas in the exposure area that indicate gradually decreasing effluent concentrations.

During the Initial, Periodic and Focussed phases of the monitoring program, an effect on the benthic invertebrate community will be determined if there are statistical differences in any of the following key descriptors :

1. **Total invertebrate density**
2. **Taxon (i.e., Family) Richness**
3. **Simpson's Diversity Index**
4. **Bray-Curtis Index**

For the benthic invertebrate component it is recommended that the following community descriptors also be calculated and reported:

Evenness
Taxon (i.e., Family) density
Taxon (i.e., Family) proportion
Taxon (i.e., Family) presence/absence

These descriptors are largely summary metrics selected to encompass the range of effects, which may be a result of mine effluent. For a more detailed discussion of these descriptors refer to section 5.18.

Evenness, Taxon density, proportion or presence/absence **above** are **included** for **the interpretation of effects, where appropriate,** in the Initial, Periodic or Focussed Monitoring or for **Investigation of Cause**. Although these are the recommended descriptors for determining and interpreting effects, there may be other site-specific descriptors, which are useful for the interpretation of monitoring data (see Resh et al. 1995 for a review).

For the statistical analyses, and determination of sufficient power, the recommendations for setting of effect size, α, and β discussed in section 5.11.2 are also applicable. The recommendation in this previous section was to set α, and βequal, at 0.10 or less. The appropriate method of analysis for each of the five study design options (e.g., ANOVA, ANCOVA, regression, multivariate analysis) is indicated in section 5.6.

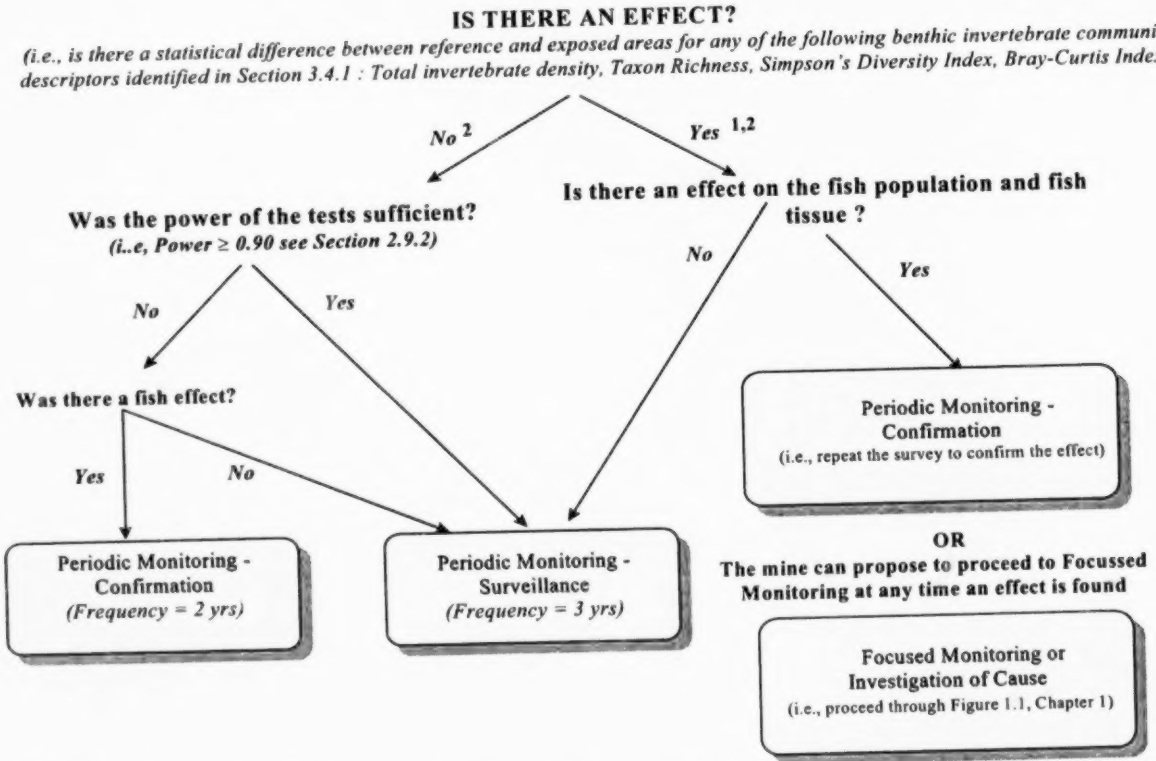
A final caveat is in order regarding effects on the benthic invertebrate community. As an effect on the benthic invertebrate community is defined as a statistical difference between benthic invertebrate community measurements taken in an exposure area and reference area of a statistical difference between sampling areas in the exposure area that indicate gradually decreasing effluent concentrations, it is essential for the mine to select a site specific study design which will allow an appropriate evaluation. Critical to the study design is the selection of an appropriate reference area or areas (discussed in section 5.8). The importance of proper reference area selection is underscored by the following, potentially frequent mining example. If a mine performs a simple control/impact design with the reference area placed upstream, then differences between upstream and downstream communities will be those determining the presence or absence of effects. However, if the downstream benthic communities are modified due to a factor such as the restoration of an upstream flow disruption (e.g., from a dam), then these communities, although different from upstream communities, may be more similar to (but perhaps not exactly the same) as communities at a reference area chosen in a drainage basin adjacent to (or even further afield than) the mine drainage basin. In this example, selection of an additional reference area (see Figure 5.4a for an example) may well be worth the extra cost involved so that site specific interpretation and the appropriate assessment of effects could be accomplished (this example of significant upstream-downstream differences may not necessarily be considered an effect if sufficient additional evidence suggests otherwise).

5.19.2 What is the next MM EMM Step?

Once the monitoring data has been analyzed decisions regarding the next step in the EEM program must be made (Figure 5.11). The next step in the monitoring program is dependent on the relationship between several key factors, which are briefly discussed below.

Figure 5-10: Decision Trees for Data Assessment

A) Recommended Monitoring Frequencies: First Benthic Invertebrate Survey Results
(or acceptable historical data Sections 2.1 and 3.4.1)


IS THERE AN EFFECT?
(i.e., is there a statistical difference between reference and exposed areas for any of the following benthic invertebrate community descriptors identified in Section 3.4.1 : Total invertebrate density, Taxon Richness, Simpson's Diversity Index, Bray-Curtis Index)
No 2
Yes 1,2
Was the power of the tests sufficient?
(i..e, Power ≥ 0.90 see Section 2.9.2)
Is there an effect on the fish population and fish tissue ?
No
Yes
No
Yes
Was there a fish effect?
Yes
No
Periodic Monitoring - Confirmation
(Frequency = 2 yrs)
Periodic Monitoring - Surveillance
(Frequency = 3 yrs)
Periodic Monitoring - Confirmation
(i.e., repeat the survey to confirm the effect)
OR
The mine can propose to proceed to Focussed Monitoring at any time an effect is found
Focused Monitoring or Investigation of Cause
(i.e., proceed through Figure 1.1, Chapter 1)

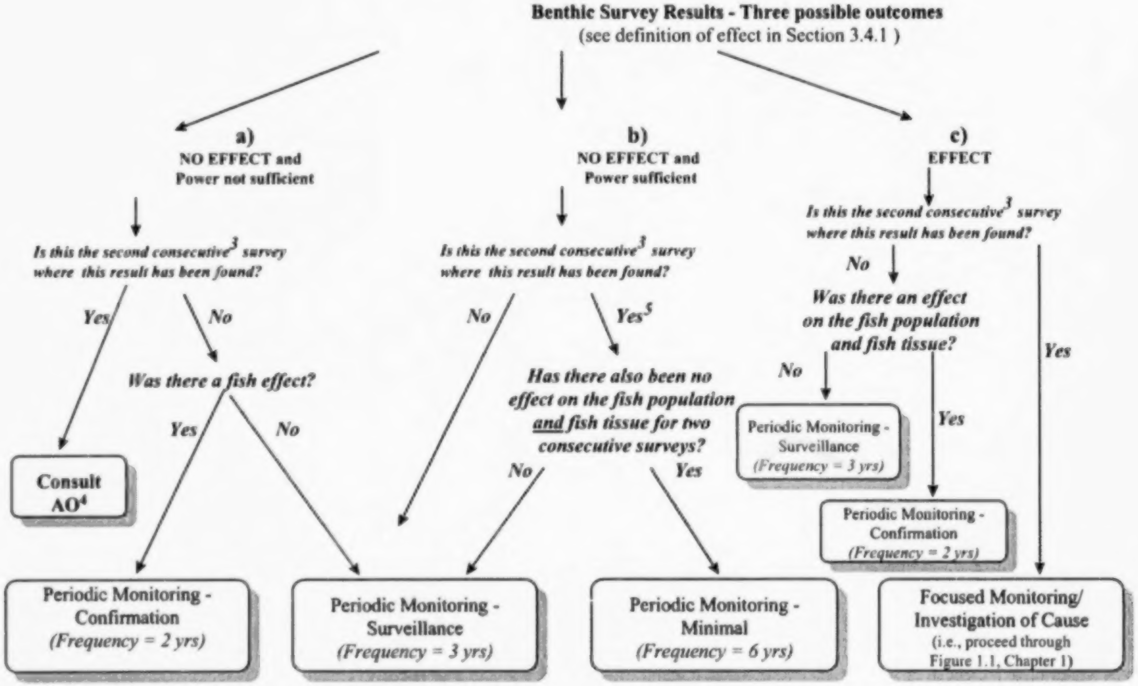
1 If there is a known effect of the mine effluent on the receiving environment but there is insufficient data to perform statistical analyses, the mine can answer **YES** to this question and proceed to Option 1 or 2.

See section 5.19.1 regarding interpretation of effects.

Figure 5.11 (con't): Decision Trees for Data Assessment

B) Recommended Monitoring Frequencies: Second or Subsequent Benthic Invertebrate Survey Results

(or acceptable historical data, Sections 2.1 and 3.4.1)


Benthic Survey Results - Three possible outcomes
(see definition of effect in Section 3.4.1)
a) NO EFFECT and Power not sufficient
Is this the second consecutive3 survey where this result has been found?
Yes
No
Was there a fish effect?
Yes
No
Consult AO4
Periodic Monitoring - Confirmation (Frequency = 2 yrs)
b) NO EFFECT and Power sufficient
Is this the second consecutive3 survey where this result has been found?
No
Yes5
Has there also been no effect on the fish population and fish tissue for two consecutive surveys?
No
Yes
Periodic Monitoring - Surveillance (Frequency = 3 yrs)
Periodic Monitoring - Minimal (Frequency = 6 yrs)
c) EFFECT
Is this the second consecutive3 survey where this result has been found?
No
Was there an effect on the fish population and fish tissue?
No
Yes
Periodic Monitoring - Surveillance (Frequency = 3 yrs)
Periodic Monitoring - Confirmation (Frequency = 2 yrs)
Yes
Focused Monitoring/ Investigation of Cause (i.e., proceed through Figure 1.1, Chapter 1)

3 Consecutive surveys do not have a set frequency, but 6 years should be considered the maximum allowable time between consecutive surveys.
4 Before another redesign exercise is implemented discussions with the AO are recommended regarding why sufficient power is not being attained. If another redesign of the field monitoring program is unlikely to solve the power issues, consideration of an alternative method may be appropriate (see Table 5.2).
5 Use of historical data at this point is not recommended.

The factors that may contribute to the decision making are highlighted below.

1. The statistical outcome of the benthic invertebrate survey.

There are three possible statistical outcomes from the benthic invertebrate survey;

a) no effect is seen but power is not sufficient (i.e., power < 0.90),
b) no effect is seen and power is sufficient (i.e., power≥or 0.90), or
c) an effect is seen.

If any of the key descriptors (total invertebrate density, taxon richness, Simpson's diversity index and Bray-Curtis index) demonstrate a statistical difference between exposure and reference areas (or the gradient), then the conclusion is that there is an effect on the benthic invertebrate community. This result can be obtained by various statistical methods and those applied depend on the study design of the monitoring program.

If no effect is seen, but there was insufficient power, it is recommended that mines consider this an effect on the benthic invertebrate community and proceed with the monitoring frequencies that would be necessary, based on the results of fish monitoring. For example, if there was no effect on the benthic invertebrate community with insufficient power, and there was an effect on the fish population and fish tissue in the first biological monitoring study, it would be recommended that the next biological monitoring study be completed within 24 months (Periodic Monitoring - confirmation). If the power was insufficient, the mine may reconsider the number of sampling stations or the sampling design that was used, in order to design a study with sufficient power in the next survey.

2. The results from the previous fish survey (see Chapter 4 for more detail).

There are four possible outcomes from the fish survey (fish population and fish tissue), which will be considered during the determination of the next monitoring phase for biological monitoring;

- A fish effect is seen in one survey
- A fish effect is seen in two or more consecutive surveys
- No fish effect is seen in one survey

- No fish effect is seen in two or more consecutive surveys

3. Changes in the environmental conditions or in the mine operating conditions

Effluent characterization and water quality monitoring continues for the same parameters and at the same frequency level throughout the EEM program (See Chapter 6 for details). This information can be utilized to detect whether any changes are occurring in either the mine operating conditions (i.e., effluent characteristics) or the natural environmental conditions (in either the reference or exposure areas). These changes may be relevant if the mine is in the Periodic Monitoring - Minimal frequency.

It is recommended that the mine complete a biological monitoring study within 24 months if either the environmental conditions or the mine operating conditions change (e.g., a change resulting in an increased loading of deleterious substances to the receiving environment), which may cause an increased risk of impacts to the benthic invertebrate community, fish or fish tissue. If, however, the environmental conditions change as a result of anthropogenic factors not related to the mine effluent, the mine may wish to reconsider the monitoring strategy that was previously used. Furthermore, if a mine is in the Minimal Monitoring phase, in addition to the specific effluent and water quality data discussed above, a careful review of all available data currently being collected should occur on an ongoing basis (e.g., yearly) and be taken into consideration for decisions regarding monitoring frequencies.

4. If statistical testing is not available from historical data

At the mine's discretion, if there is a known effect on the benthic invertebrate community from the mine effluent but there is insufficient data to perform statistical analyses, the mine can propose to answer **YES** to the question regarding whether there is an effect. An example of this situation would be if a mine has many years of historical monitoring data but the previous study designs did not include replicate stations at the scale described in this guidance document (e.g., replication may only have occurred at the field sub-sample level). However, in this situation, the many years of data may indicate a reasonable probability of an effect. A weight-of-evidence approach (section 5.23.1) may be applicable in this case to allow effective use of the historic monitoring data. In this kind of situation, a mine may proceed to Focussed Monitoring or Investigation of Cause.

5.19.3 Summary of biological monitoring frequencies for the benthic invertebrate community after data has been assessed

Depending on whether or not there are any statistical differences in the measurements that are used to determine effects in the benthic invertebrate community detected between reference and exposure areas (or the gradient), then the EEM program will move to one of five options (Figure 5.11). The power of the benthic invertebrate community survey will be taken into consideration and the results of the fish survey and fish tissue analysis will be used in the decision making process.

1) **Periodic Monitoring - Surveillance**: effects in one or two (but not all three) of the benthic invertebrate community, fish population or fish tissue. Monitoring, including submission of the interpretative report, must be completed within 36 months of the submission of the previous interpretative report *(Schedule 5, S. 22(1))*.
2) **Periodic Monitoring – Confirmation**: effects in **all three** of the benthic invertebrate community, fish population and fish tissue. Monitoring, including submission of the interpretative report, completed within 24 months of the submission of the previous interpretative report *(Schedule 5, S. 22(2)(a))*. If no effect was found in the benthic invertebrate community, but there was insufficient power, and there were effects in the fish population and fish tissue, it is recommended the mine proceed with this monitoring frequency.
3) **Periodic Monitoring - Minimal**: no effects in two consecutive surveys in the benthic invertebrate community, fish population **and** fish tissue. Monitoring, including submission of the interpretative report, completed within 72 months of the submission of the previous interpretative report *(Schedule 5, S. 22(2)(b))*.
4) **Focussed Monitoring**: effects in two consecutive surveys in the benthic invertebrate community, fish population **or** fish tissue. Monitoring to determine the magnitude and geographic extent, including submission of the interpretative report, completed within 24 months of the submission of the previous interpretative report *(Schedule 5, S. 22(2)(c))*
5) **Investigation of Cause**: magnitude and geographic extent of effects known, but the cause is not known. Investigation of Cause study, including submission of the interpretative report, completed within 24 months of the submission of the previous interpretative report *(Schedule 5, S. 22(2)(c))*.

5.20 EEM program options after an effect has been established

If an effect on the benthic invertebrate community is found, the mine will then proceed through the overall EEM program decision tree (Figure 1.1). The next question to be addressed is:

5.20.1 Is the effect mine related?

An assessment of whether the effect is mine related could include asking the following questions:

- Is the cause of the effect known or suspected?
- Can the effect be related to a natural change in the aquatic-receiving

environment?

- Can the effect be reasonably correlated to an anthropogenic cause, other than the mine effluent?
- Is there a weight-of-evidence approach which can indicate a causal link (see Section 5.23.1)?

This series of questions is provided as examples for the type of approach, which may allow the determination of whether the observed effect is mine related. If it is questionable whether or not the effect is mine related after one monitoring phase due to the presence of confounding factors, the mine should reconsider the study design for Periodic Monitoring - confirmation, in order to ensure that the effect is mine related. If the effect has been confirmed, and the cause of the effect is unknown, then the mine should proceed to the next step in Data Assessment and Interpretation, determining the magnitude and geographic extent of the effect.

5.20.2 Is the magnitude and geographical extent known?

The objective of Focussed Monitoring is to conduct an assessment of the magnitude and geographical extent of the effect. Focussed Monitoring may include the following questions:

Magnitude:

1) How many taxonomic groups are affected?
2) What is the magnitude (e.g., the amount of change in abundance) of the effect on the taxonomic groups affected?
3) Is there an effect in other benthic invertebrate community members, such as periphyton and macrophytes, which are present in the reference area and would be expected to be present in the exposure area?

Geographical Extent

1) What is the geographical area affected?
2) Are the benthic invertebrate communities at the sampling stations furthest from the effluent discharge similar to reference conditions?

5.20.3 Is the cause of the effect known?

The objective of Investigation of Cause is to conduct an assessment of the cause of the effect and may include the following questions:

1) What are the possible mine-related causal agents?
2) Is there a correlation between the type of effect observed on the benthic invertebrate community and a mine-related causal agent?

3) Is there a weight-of-evidence approach which can indicate a causal link (see Section 5.23.1

5.21 Methods for Benthic Invertebrate Monitoring

This section outlines recommended field methods for the sampling of benthic invertebrate communities and the processing of samples in the laboratory and refers to literature, which discuss these methods in more detail.

5.21.1 Field Methods

5.21.1.1 Sampler Mesh Sizes

Benthic samples typically contain varying amounts of fine sediment and debris. To expedite transfer to sample containers, storage, and shipping, these samples should be reduced in the field by sieving. Field sieving should be done wherever possible, immediately after sample retrieval and before preservation, as many organisms become fragile and brittle after preservation. Various techniques for sieving are available, but most involve washing the sample with a sieve or sieve bucket device.

The recommendation for sieve and/or mesh size for all freshwater mines is 500 µm.

In freshwater, macroinvertebrates are defined as those retained by mesh sizes of 200-500 µm (Slack et al. 1973; Weber 1973; Wiederholm 1980; Suess 1982), although immature life stages of some taxa may be smaller and some adult life stages may be larger.

Note that these mesh size are applicable to all equipment used in the field and laboratory (i.e., both the Nitex mesh on the benthic samplers and sieving apparatus).

However, in some site-specific circumstances it may be desirable for the field samples to be screened for smaller organisms by using a smaller sieve size. Some examples of situations where the the use of a smaller mesh size (less than 500 µm) may be appropriate include the following:

1) for comparative purposes if historical benthic surveys for the system under investigation utilized smaller mesh sizes, or
2) if sampling needs to be conducted, for logistic reasons, at times when organisms are very small , (Rees (1984), Barber and Kevern (1974) and Jonasson (1955) provide information on seasonal effects of mesh size)

In these aforementioned cases, it is highly recommended that a stack of screens be used which minimally have the mandatory sieve sizes and then any other smaller sizes which are appropriate. This procedure simultaneously allows site-specific concerns to be addressed and fulfills EEM objectives by allowing for national or regional comparisons to be conducted on the standardized mesh sizes. Note, that sieving with the finest scale sieve can be done in the field, as long as the appropriate fractionation of the sample is performed in the laboratory before processing.

For marine organisms, samples should be sieved with seawater rather than freshwater, since the osmotic shock of freshwater may cause cell bursting and gross distortion of the animals and where appropriate, field water used to sieve should be screened for ambient organisms with a mesh smaller than the required minimum screen size used for the study. In addition, extreme care should be taken during washing of samples to avoid breakage of specimens, which can greatly reduce taxonomic efficiency and cost-effectiveness. Methods have been described to reduce breakage, particularly in marine samples (Gray et al. 1991).

In marine systems, it is recommended a **stacked set of 1.0 and 0.5 mm** screens in the field with the 0.5 mm samples being archived and processed only if appropriate. Marine macrobenthos are typically those retained by sieves with 500-1000 µm mesh (Reish 1959, Thiel 1975, Pearson 1975, Holme and McIntyre 1984, Gray et al. 1991). It is estimated that a 1.0 mm sieve will retain about 95% of the biomass of marine macrofauna (Reish 1959) while reducing the numbers of juvenile taxa and meiofauna present in samples which respond functionally differently to environmental perturbation than adult macrofauna (Schwinghamer 1981, 1983; Warwick 1986).

Studying smaller benthic organisms for Focussed Monitoring in marine systems may include assessment of meiofauna such as nematodes, copepods and smaller oligochaetes; or it may include assessment of living and dead foraminifera (Schwinghamer 1981, 1983); or it may include more detailed assessment of juvenile forms of macrofauna. All of these approaches require the use of smaller mesh sizes and/or different samplers (cores may be more appropriate than grabs – see Holme and McIntyre 1984) than are currently recommended above. However, if smaller forms are important, simply adding an additional sieve may not fulfill this function. The appropriateness of the sampling techniques should be assessed for smaller forms. For marine environments, Gray et al. (1991) noted that collection of meiobenthos is most appropriately performed with corer samplers which are not recommended sampling devices for the EEM program. Thus, before simply screening for smaller organisms, appropriate protocols should be implemented.

5.21.1.2 Sampling Equipment

During the design phase of the survey, certain decisions need to be made concerning the selection of specific sampling equipment. The two major considerations for benthic surveys are mesh size (see previous section) and quantitative sampling equipment. Quantitative sampling of benthic communities is carried out using devices that sample a known area or volume of habitat, such as grab samplers or stream net samplers. Each sampling device should be non-selective and suitable for a particular condition of substrate. Benthic samples collected from natural substrates provide an indication of historical and current stressors. Therefore, samplers that collect benthos communities from the bottom sediments are recommended unless this is not possible due to physical constraints. Samplers are to be consistent within a habitat class between all stations and areas. However, different samplers may be used in the same survey if they are used to sample different habitat classes. For example, if both erosional and depositional habitat classes are sampled, it would be reasonable to use one of the recommended grabs for the depositional habitat, but a Hess-type sampler, for the erosional habitat. It is recommended that grab samplers with screens on the top and top-opening gates be used so that the bow-wave ahead of them is reduced and less substrate is lost, and for examination (and possibly sediment chemistry analysis) of undisturbed surface layer in sediment samples.

Standardization of benthic samplers is necessary for regional or national comparisons of benthic invertebrate survey data to be valid. Specific recommendations are forwarded as to the appropriate samplers for the various habitat classes encountered during EEM benthic surveys. These options are discussed below as they pertain to each potential receiving environment and summaries of the advantages and disadvantages of the recommended samplers are presented in Table 5.8. The selection of a sampler may also be influenced by the type of sampler(s) used in previous surveys of that system. To ensure that surveys can be compared with previous historical surveys, it would be useful to use similar sampling equipment. For more detailed information please refer to the bibliographies on quantitative samplers and appropriate methodologies which have been prepared by Klemm et al. (1990), Eleftherious and Holme (1984), Elliott and Tullett (1978,1983) and Rosenberg (1978), Downing (1984) and Mason (1991a). See also Rabeni and Gibbs (1978) and Alberta Environment (1990).

Standardization of techniques does not only apply to the actual samplers but to the level of expertise required to correctly deploy the sampler. Employees using the sampler must be properly trained so that any areas where subjectivity may be interjected is limited. For example, when sampling erosional zones in rivers, the depth to which the substrate is disturbed within the net-sampler should be standardized since some individuals may be more enthusiastic in regards to "stirring" up the substrate than others. It is essential that the study leader be well versed in benthic invertebrate sampling and conduct effective training sessions with the employees who will be performing the field sampling. If training is done effectively, operator error can be eliminated (Reynoldson and Rosenberg 1996).

The use of artificial substrates for benthic invertebrate collections is generally not recommended as a sampling protocol in EEM.

There is no advantage to be gained from using artificial substrates where conventional sampling techniques provide at least as reliable data without many drawbacks and difficulties of artificial substrates (AETE, 1995). Artificial substrates do not collect a representative sample of the indigenous benthic invertebrate community at the site where they are placed, but rather select for mobile, drift-prone species of hard substrata. In addition, artificial substrates do not effectively monitor the effects of sediments or sediment-bound contaminants on aquatic biota because sediment dwelling taxa tend to be under-represented in artificial substrate samples and finally, the invertebrate community represented by artificial substrates indicates conditions during the period of exposure only and does not integrate long-term effects. Therefore, the use of artificial substrates for benthic invertebrate collections may fail to indicate the effects from effluents, particularly where non-mobile species, sediment bound contaminants, or longer term integration of effects are important. However, it is recognized that there may be a limited number of cases where there is either a long history of artificial substrate use in a particular ecosystem, or extreme habitat conditions (e.g., very deep, fast water systems) where the use of artificial substrates is the only feasible *field* method available. In these cases the use of artificial substrates may be considered along with other alternatives (Table 5.2) provided this method can determine if there are effects on the benthic invertebrate community in a scientifically defensible manner.

Depositional Habitats - Freshwater

Grab samplers, sometimes called dredges, are devices with spring-loaded or gravity-activated jaws that "bite" into unconsolidated substrates (sand, silt, mud, etc.) to enclose a defined surface area of the bottom. These devices are generally lowered on a line or cable from a survey vessel to the bottom, sometimes with the aid of a winch. The effectiveness of these samplers depends largely on the relationship between the design of the sampler and the type of substrate. If the sampler type chosen is not suited to the substratum present, it can affect sampling efficiency. Factors which may be affected include, the depth of penetration, completeness of the closure of the jaws, and the subsequent loss of material during retrieval. In depositional zones in freshwater rivers or lakes, Ponar or Ekman grabs are suggested as standard samplers for EEM invertebrate surveys (Table 5.8).

Erosional Habitats - Freshwater

Stream-Net Samplers are devices used for collecting benthic invertebrates in erosional riverine environments. They use mesh of various sizes (but see Section 5.21.1.1 for discussion on mesh size) to sieve organisms from water flowing through the mesh after disturbance of a known area of the substrate. It is recommended that erosional habitats in freshwater environments be sampled with the Neill-Hess cylinder type samplers which allow unit area (typically, $0.1m^2$) estimates to be made. However, one draw back to cylinder samplers in streams is an incompatibility with size of substrate. For example, in some large rivers, mean particle size may be too great for the Neill-Hess cylinder to effectively sample the benthic invertebrates. In such cases, a recently tested U-net sampler (Scrimgeour et al. 1993) can provide area-limited samples and can be adjusted accordingly to the size of the substrate. This sampler has been used successfully for a range of substrate sizes (Glozier 1989) and can sample either individual stones or a delimited area. Kick net samplers, although they do not provide an area delimited estimate, have been used widely in the United Kingdom, the U.S.A., Australia and Canada for large scale monitoring programs (Reynoldson et al. 1995). Kick sampling is particularly appropriate for the Reference Condition Approach where many stations are sampled and a timed kick sample is taken at each station to estimate benthic community descriptors. Standardization of kick sampling techniques is essential for comparative purposes and can be obtained with minimal training (Reynoldson and Rosenberg 1996). The kick sampling method involves a single composite sample collected at each station by a 3 min traveling kick method (Reynoldson et.al.,1997). Note that separately preserved field sub-samples are not required for the kick sampling technique recommended for RCA.

For difficult habitats (e.g., very deep, slow flowing areas or areas with hard substrates) there may be alternative samplers available such as the metal quadrat and air lift system. However, for national or regional comparative purposes the list of recommended samplers should be sufficient for sampling the majority of ecologically relevant habitats. If habitats are extremely difficult to sample, alternative approaches may be considered (Section 5.23).

Table 5-8a: Summary criteria for recommended samplers in freshwater

Sampler	Characteristics
Freshwater - *Erosional Habitats in Rivers*	
Hess Sampler or **Neill Cylinder Sampler** **AREA : 0.1 m^2**	• **Habitats and substrates sampled:** shallow, flowing streams, with good current, rubble substrate, mud, sand, gravel • **Effectiveness:** quantitative when used by experienced biologist, performance depends on current and substrate • **Advantages:** encloses area sampled, easily transported or constructed, samples a unit area, stable platform; can be used in weed beds. • **Limitations:** difficult to set in some substrate types, that is large rubble; cannot be used efficiently in still, slow flowing streams.
U-net samplers **AREA : 0.06 m^2 but any size can be specified during construction**	• **Habitats and substrates sampled:** shallow, flowing streams, with good current, rubble substrate, mud, sand, gravel, and large cobble • **Effectiveness:** quantitative when used by experienced biologist, performance depends on current and substrate • **Advantages:** easily transported or constructed, samples a unit area, stable platform; size of frame can be constructed to adapt to various substrate types. • **Limitations:** cannot be used efficiently in still, slow flowing streams.
Kick net samplers **AREA : no set area but D-frame net is typically 30 cm wide**	• **Habitats and substrates sampled**: shallow, flowing streams, with good current, rubble substrate, mud, sand, gravel, and large cobble • **Effectiveness**: performance depends on current and substrate • **Advantages:** use easily learned, easily transported and constructed, • **Limitations**: not truly quantitative, does not sample a unit area, standardization set by controlling the amount of time kicking
Freshwater - *Depositional Habitats in Rivers and Lakes -*	
Standard Ponar Grab **AREA : 0.05 m^2**	• **Habitats and substrates sampled:** freshwater lakes, rivers, estuaries, and reservoirs with hard and soft sediments such as clay, hard pan, sand, and muck; somewhat less efficient in softer sediments. • **Effectiveness:** not entirely adequate for deep burrowing organisms in soft sediments; very efficient for coarse sediments; collects both qualitative and quantitative samples • **Advantages:** better penetration than other grabs; side plates and top screens reduce washout, shock waves and substrate disturbance; best quantitative grab sampler for freshwater use. • **Limitations:** a heavy grab that requires use of a boat with winch and cable; stones, pebbles, and other debris can hold jaws open causing loss of sample.
EkmanGrab (Standard, Large and Extra-Large) **AREA : 0.02 m^{2,} 0.05 m^2 and 0.09 m^2,**	• **Habitats and substrates sampled:** freshwater rivers, lakes, and reservoirs where there is little current; soft sediments, e.g., muck and silt. • **Effectiveness:** efficient only in soft sediments but weights can be added for deeper penetration in fine sand; collects both qualitative and quantitative samples. • **Advantages:** easy to operate by hand, can be pushed into substrate in shallow water; hinged doors at top reduce washout, shock waves, and disturbance of the substrate; comes in a range of sizes. • **Limitations:** light weight so that jaw will not penetrate hard substrates; jaws often do not close completely due to blocking of jaws or failure of closing mechanism; inefficient in deep water or where current is moderate.

Source: Eleftherious and Holme (1984), Klemm et al. (1990), Scrimgeour et al. (1993)

Depositional Habitats -Marine/Estuarine -

Depositional habitats in marine environments can be sampled with the Smith-McIntyre grab, a modified Van Veen grab which are suitable and available in Canada. A good review of marine sampling methods is available in Eleftheriou and Holme (1984). However, in shallow subtidal areas where there is not enough water depth to allow the deployment of the larger grabs, the use of a smaller, or mini (petite)-Ponar grab can be substituted. This grab is deployable from small inflatable boats and can be retrieved by hand.

Intertidal soft substrates may be sampled using any device which demarcates at least $0.1 m^2$ area. The soft substrate is then removed to a standard depth of 10 cm using whatever device is appropriate. Note that in general, the lowest intertidal level available for sampling is preferred because less harsh physical conditions promote higher species richness and abundance.

Erosional Habitats-Marine/Estuarine

In marine/estuarine environments, large unconsolidated sediments such as gravel may be sampled with grab samplers. If not, hard substrates in erosional habitats (intertidal and subtidal) should be sampled using quadrats with a minimum size of $0.1 m^{2.}$ However, some other quantitative techniques may be recommended to collect marine shellfish and other large species. These may include hand collections by divers, remote photographic techniques from defined surface areas (Eleftheriou and Holme 1984, Gray et al. 1991), and collections from defined boundaries along transects. An outline of a marine sampling protocol is described in a series of Puget Sound Estuary Program Reports (Tetra Tech 1986a, b, 1987). When done properly, photographic surveys can be quantitative, at least for larger epibenthic organisms (c.f. Burd and Brinkhurst 1984). Processing costs tend to be considerably less than for soft-bottom surveys using grabs or cores.

The intertidal zone should be sampled if the effluent plume impinges substantially on it. Determining the tidal level of greatest interest for examining mine impacts will involve logistical considerations. Basically, the lower in the intertidal the surveys can be conducted, the better, since less harsh conditions create less patchiness and higher diversity in flora and fauna (inter-sample variability). Coastal plants and animals in this habitat typically exhibit vertical distributions that reflect gradients in environmental parameters such as air exposure, temperature (including freezing), salinity, light intensity and daylength, abrasion due to logs or ice and wave shock. These gradients should be considered in planning and undertaking biological surveys in the intertidal environment. Sampling protocols for this area will be somewhat different from those described in the earlier sections (for review see Gray et. al. 1991). Wherever possible, semi-quantitative surveys using quadrat areas of $0.1 m^2$ should be done. Determining the substrate or habitat type to be sampled is dependent on sampling limitations and the dominant habitat present (see Section 5.19 for discussion on dominant habitat selection) However, if it is not feasible or ecologically sound to collect samples, then visual surveys are recommended. If approved, a visual survey would include approaches such as recording and mapping (at a gross scale 5,000:1) the major biological features for assessment of gross changes in the biological community.

Table 5-8B: Summary criteria for recommended samplers in marine/estuarine habitat

Sampler	Characteristics
Marine/ Estuarine - Erosional Habitats-	
Quadrat samplers **AREA : 0.1 m^2 frames**	• **Habitats and substrates sampled:** large gravel or rocky bottom intertidal and subtidal (samples may be collected, photographed, mapped etc.). • **Effectiveness:** depends on visual skills of field person, sampling often limited to larger organisms by field personnel. • **Advantages:** For subtidal surveys, SCUBA transects with quadrats for counting megafauna are the only semi-quantitative method available. When depth are too deep for SCUBA, an ROV camera may replace the diver. • **Limitations:** for visual surveys, smaller organisms are underrepresented, can be logistically difficult in subtidal areas with more work and time in the filed than samples which can be brought back to the lab, semi-quantitative only.
Marine/ Estuarine - Depositional Habitats -	
Petite Ponar Grab **AREA : 0.02 m^2**	• **Habitats and substrates sampled:** freshwater lakes, rivers, and reservoirs and estuaries with moderately hard sediments such as sand, silt, and mud; will not penetrate clay; somewhat less efficient in soft sediments and coarse gravel. • **Effectiveness:** limited depth penetration; not useful in clay because of light weight, half recommended sample area size requires twice as many grabs. • **Advantages:** good penetration for such a small grab; can be operated by hand. • **Limitations:** jaws can be blocked by stones, sticks and other debris causing loss of part of the sample; not efficient in swiftly flowing water; smaller size makes it inefficient for sampling larger benthic organisms; not entirely adequate for deep burrowing organism in soft sediments
Standard Ponar Grab **AREA : 0.05 m^2**	• **Habitats and substrates sampled:** freshwater lakes, rivers, estuaries, and reservoirs with hard and soft sediments such as clay, hard pan, sand, gravel, and muck; somewhat less efficient in softer sediments. • **Effectiveness:** not entirely adequate for deep burrowing organisms in soft sediments; very efficient for coarse sediments; collects both qualitative and quantitative samples • **Advantages:** better penetration than other grabs; side plates and top screens reduce washout, shock waves and substrate disturbance; best quantitative grab sampler for freshwater use. • **Limitations:** a heavy grab that requires use of a boat with winch and cable; stones, pebbles, and other debris can hold jaws open causing loss of sample.
Smith-McIntyre Grab **AREA : 0.1 m^2**	• **Habitats and substrates sampled:** marine and estuaries; adaptable to large rivers, lakes, and reservoirs with sand, gravel, clay, and similar substrates. • **Effectiveness:** has been widely used for sampling in marine and estuarine habitats, penetration - 17 mm in silt and clay, 5 dm in coarse sand. • **Advantages:** provides reasonably quantitative samples; trigger plates help penetrate the substrate, top screens and stabilizing frame reduces washout and ensures upright landing. • **Limitations:** very heavy, needs boat and power winch; spring-loaded jaws could be hazardous; inefficient for collecting deep burrowing organisms; jaws can be blocked by debris.
Van Veen Grab **AREA : 0.1 or 0.2 m^2**	• **Habitats and substrates sampled:** marine and estuaries with sand, gravel, mud, clay, and similar substrates; could be adapted to freshwater. • **Effectiveness:** penetrates to a depth of 5 to 7 cm. • **Advantages:** jaws close better than the Petersen grab; samples most types of sediments; comes in a range of sizes. • **Limitations:** a very heavy grab that requires a large boat and power winch; jaws may become blocked by debris such as rocks and sticks; not useful for deep burrowing organisms.
Quadrat to depth of 10 cm **AREA : 0.1 m^2**	• **Habitats and substrates sampled:** intertidal marine/estuarine soft substrates to fine gravel. • **Effectiveness:** very effective with care of field personnel. • **Advantages:** inexpensive, allows sampling of substrates not accessible with grab samples • **Limitations:** may not sample deep burrowing organisms in sandy sediments.

Source: Eleftherious and Holme (1984), Klemm et al. (1990)

5.21.1.3 Specimen Preservation

All samples should be fixed in the field in 10% buffered formalin to prevent damage to freshwater and marine worms and polychaetes. Formalin is also important for the proper preservation of most aquatic insects. Preservation directly in ethanol often results in soft, difficult to handle specimens (but see laboratory options for transferring to ethanol at a later stage). After preservation in the field, samples should be gently mixed several times to ensure that the preservative has thoroughly penetrated any fine material which may be present in the sample. Because formalin is a carcinogen and an irritant to workers, gloves and eye protection are needed to protect against splashes, and should be considered mandatory safety equipment. Furthermore, non-breakable sample jars should be sealed with parafilm, doubled bagged for transport back to the laboratory facilities and adequately labeled.

5.21.2 Methods for Supporting Environmental Variables

Supporting environmental variables are recommended to be measured at the time of the benthic survey. Exposure variables are those variables measured as part of water quality monitoring (Chapter 6). Chapter 6 also lists other water quality variables that may be useful to monitor on a site-specific basis. It is recommended that explanatory variables listed in Table 5.5 be measured quantitatively at each replicate station and the majority of methods are included in either Chapter 6 or 8. For those variables specific to the benthic invertebrate sampling program and not included elsewhere in the guidance document brief explanations of methods are provided below.

Wetted Width

Is the width of the wetted stream bed, recorded in meters. The measuring tape should be stretched across the width of the stream, level and perpendicular to the flow. Stream sections measured for wetted width should not include small islands or dry areas in the stream bed. For highly braided streams the wetted widths of the separate channels should be measured individually and summed for a total wetted width.

Bankfull Width

Is the width of the channel under bankfull or channel maintenance flow conditions. As stream surveys are rarely done during bankfull conditions, the width of the bankfull channel is estimated by determining where the tops of the banks are and measuring the width at that point. The bankfull stage usually coincides with the level of roots of perennial vegetation on the stream bank (i.e., bankfull is slightly above this point). As with wetted width, the measuring tape should be stretched across this width of the stream and be level and perpendicular to the flow (see Newbury and Gaboury 1993 for further descriptions).

Embeddedness

Embeddedness is a supporting variable measured in riffle habitats. In undisturbed streams, fine sediments do not generally accumulate in large quantities in riffle gravel/cobble substratum. To measure the percent embeddedness with a station, 5-10 cobbles are examined to determine the percent which are embedded in fine sediments. Avoid rocks which have been recently overturned by high flows or benthic sampling procedures. After extracting a cobble from the stream bed it is examined to determine the percent of the rock which was buried in the fine sediments. A stain line is often present on the rock indicating the level of burial (see US EPA 1995 for further descriptions).

Riparian Vegetation and Canopy Cover

Qualitative notes on riparian vegetation should be taken. In addition, the width of the riparian zone is an indicator of the extent to which the zone has been disturbed. The width of the undisturbed riparian vegetation can be measured at each station by establishing two transects, one on either bank and measuring the width in meters (see Tabacchi et al. 1998, Patten 1998 and Ladson and White 1999 for further details).

Canopy cover can be measured with a spherical densiometer. This device consists of a concave mirror with 24 – ¼" squares engraved on the surface. The standard method with this device includes taking a measure of the number of squares which are covered with vegetation while the operator is standing in the center of a riffle and viewing the mirror appropriately. Four estimates are taken in each station while standing in the middle of the channel, one facing in each direction (N,W,S,E). Detailed methods are described in Lemmon (1957) and included with the device.

5.21.3 QA/QC for Benthic Invertebrate Field Operations

An outline of the recommended QA/QC recommendations for the field components of the benthic invertebrate community survey is presented below. Most of the following information has been taken from Environment Canada and Department of Fisheries and Oceans (1993) and Gibbons et al. (1993).

Field sampling is the first stage of data collection. QA/QC procedures for the benthic invertebrate survey should be outlined during the development of the study design and should be followed precisely in order to maintain high quality data. Field SOP should specify sampling equipment and protocols appropriate to the study and while a QA/QC plan for field sampling can have many components, some of the main procedures are listed below.

1) all personnel involved in the field sampling should have appropriate training and experience with field equipment and objectives,
2) all safety measures should be identified, understood and adhered to,
3) collection equipment should be appropriate for the specific water body and selected invertebrate group, and must be checked frequently and maintained on a regular schedule,
4) there should be some *a priori* criteria for acceptability of samples obtained and clear directions if acceptability guidelines fail (i.e. when to retake a sample, for example, grab sample penetrations of 10 cm depth would be considered an acceptable sample, Gray et al. 1991),
5) a visual description of benthic grab samples should be recorded to describe sediment color, odour, texture and debris,
6) contamination during chemical sampling should be checked with trip blanks and equipment rinsates,
7) field sieving, if necessary, should be done as soon as possible after retrieval of samples,
8) samples should be stored in appropriate containers with appropriate preservative to prevent breakage and spoilage,
9) all sample containers should be appropriately labeled,
10) detailed field notes should be maintained in a bound notebook,
11) chain-of-custody forms and appropriate shipping and storage procedures should be used.

For further information regarding all aspects of QA/QC procedures for benthic invertebrate programs refer to the 1999 AETE report (Beak 1999).

5.22 Laboratory Methods

Please see the EEM website for further "Guidance for Sample Sorting and Subsampling Protocols for EEM Benthic Invertebrate Community Surveys" (www.ec.gc.ca\eem).

5.22.1 Sample Sorting

Upon return from the field, samples should be checked to ensure that the preservative has effectively penetrated the entire sample especially if there is large amounts of fine material. The first step in processing the invertebrate samples is to place the preserved sample into a sieve (freshwater = 500 µm) and rinse the preservative and the residual fine debris from the sample. While flushing the samples, care must be taken not to damage the organisms with too much water pressure. Samples can be sorted in water immediately, providing they are transferred back to preservative at the end of the workday (i.e., after 8 hours). Alternatively, samples may be transferred to 80% ethanol after the washing process, prior to sorting and taxonomic work. A minimum of 72 hours in formalin is recommended before transferring to ethanol. After sorting or identification, most freshwater macro-invertebrates should be stored in a solution of 70 to 80% ethanol, and 5% glycerin in vials or jars with air-tight lids. Marine samples should be transferred to 70-80% ethanol within 3 months to protect the samples from breakage as the organisms become brittle. If screw-cap lids are used, they should be sealed with wax film, and will require periodic checking (once or twice annually) to replace evaporative losses. Since preservatives can have a variety of effects on the lengths and weights of invertebrates (Ellis 1987; Leuven et al. 1985; Howmineer 1972), it is essential that the exact preservation protocol for all steps be drafted *a priori* and that it be consistently applied between all phases, areas, replicate stations and habitats. Disposal of used preservatives must be performed in accordance to provincial hazardous waste regulations (e.g., BC Special Waste Regulations, Alberta Waste Control Regulation, Ontario Regulation # 347 etc.)

For samples, which have large pieces of organic matter, the samples can be divided in the laboratory into appropriate size fractions to expedite the sorting process. The most commonly used fractions are coarse (> 1.00 mm) and fine (500 µm - 1.00 mm), which correspond to the divisions used to define coarse and fine particulate organic matter (CPOM and FPOM, respectively). If there are very large pieces or organic material or large invertebrates, it is sometimes beneficial to separate these from the rest of the sample with a 4.00-mm sieve. All fractions should then be sorted and if warranted by large numbers of organisms, the fractions can be sub-sampled independently (see Section 5.22.2). Careful note taking is recommended for these more complex sorting procedures so that densities are calculated accurately. After the initial washing and fractionation of samples, the invertebrates should be sorted from the debris by trained technicians on a gridded tray or petri dish under a dissecting microscope at 10X to 20X magnifications.

Various additional techniques exist which are designed to speed up the manual sorting required for benthic invertebrate samples. These techniques include flotation of organisms in high-density solutions (e.g. magnesium sulphate, sodium chloride, calcium chloride, sugar, D-mannitol) or bubbling to trap animals in the surface film and to promote separation from denser sediment particles. Heavier organisms such as molluscs will not float and must be removed directly with forceps. Also, various stains such as rose bengal or phloxine B at a concentration of 100 mg/l can be used to make the organisms easier to see. A cautionary note is required, however, as the efficiency of many of these techniques differ between taxa. If these techniques are used, then the QA/QC performed on these samples should be designed to detect any differential sorting efficiency of taxa caused by the staining procedures.

Proper QA/QC procedures are critical at this stage and acceptable QA/QC guidelines for the processing benthic invertebrate samples are described in section 5.22.4.

5.22.2 Laboratory Sub-sampling

Sub-sampling of the invertebrate sample in the laboratory is acceptable, providing that the method used is quantitative. Large samples or samples with large amounts of sediment debris may require laboratory sub-sampling prior to sorting. Various methods of laboratory sub-sampling are available, and are outlined in Klemm et al. (1990) as well as Canton (1991), Mason (1991b), Plafkin et al. (1989), Wrona et al. (1982), Marchant (1989) and Sebastien et al. (1988). The method used should be thoroughly documented. Generally, a minimum of one-quarter of the sample should be sorted, however, a more definitive guideline is that sub-sampling should continue until a minimum number of animals is counted to obtain a predefined variance level (see Wrona et al. 1982). For example, for the cone subsampler described by Wrona et al. (1982), as the number of animals counted exceeds 50 and approaches 100, the improvement in the error term for a given unit of additional effort decreases. The extra information gained by taking more sub-samples is not worth the extra time invested. In this case (Wrona et al. 1982) and others (Hickley 1975, Elliott 1977) the recommendation for the minimum number of animals which should be counted was 100. It is recommended that this minimum number be determined for each sorting and laboratory sub-sampling protocol utilized during an EEM program using a similar approach.

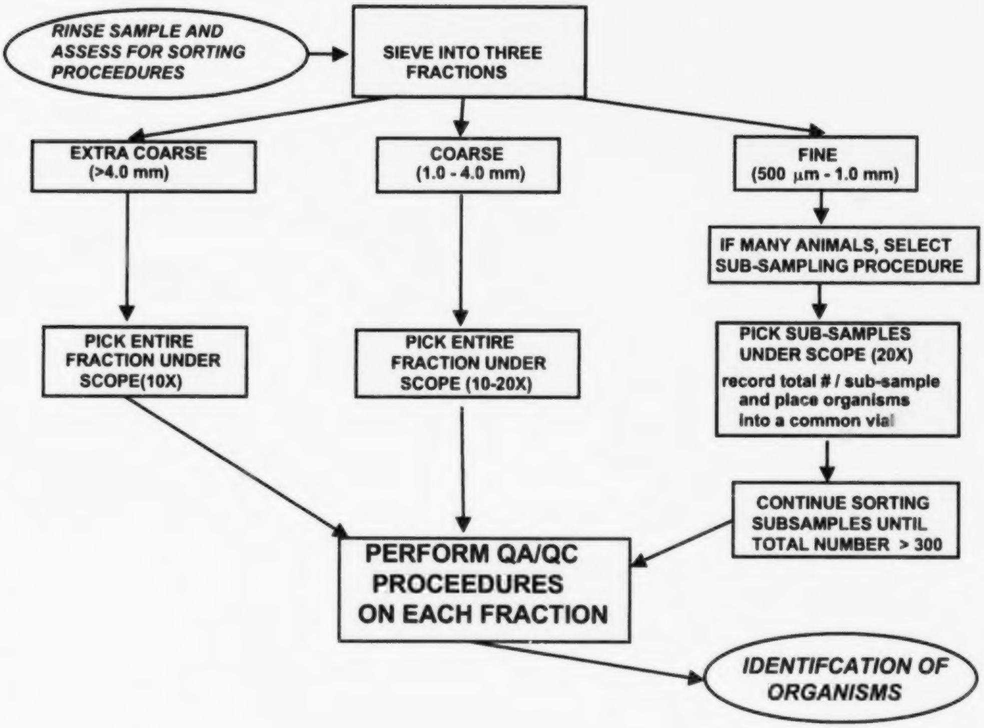
Further guidance for laboratory sorting and sub-sampling protocols include the following:

1. Sorting procedures should be reviewed after an initial examination of the type of samples. Sorting procedures may differ depending on whether large amounts of algal material, bryophyte or detrital material are present.
2. If a sample contains large numbers of single taxa (e.g. >1000 chironomids), or large numbers of organisms in a given fraction, sub-sampling can be performed for those organisms or fractions independently and the remainder of the sample sorted completely.
3. When undertaking a laboratory sub-sampling regime for a given project, error associated with sub-sampling technique should be estimated. This can be accomplished by continuing to sort sub-samples until the entire sample is sorted for a minimum of 10% of all samples. This is especially important if different, but sample appropriate, sub-sampling protocols are undertaken on samples from different areas or stations.
4. Careful note taking is recommended for these laboratory sub-sampling procedures so that the accuracy of the calculations of total density and other metrics is not compromised.

A schematic diagram of a sorting and sub-sampling protocol is provided in Figure 5.12. This is an illustrative example only which should be modified according to the type of material and the number of organisms in the sample. For example, some sub-sampling protocols are not applicable for samples with large amounts of algae or if organisms are not abundant. The example depicted in Figure 5.12 would be applicable for samples which have large pieces of detrital material (e.g., leaves or small branches) and hence the fractionation speeds the sorting procedures and large numbers of small organisms which would be concentrated into the fine fraction (500µm -1.00mm). Each laboratory should outline sample-specific sorting and sub-sampling procedures and provide full documentation of the methods used.

Figure 5-11: Benthic invertebrate sorting and sub-sampling protocol, which would be applicable for samples with detrital material and large numbers of small organisms.

This is an illustrative example only, which should be modified as necessary for site-specific samples.


RINSE SAMPLE AND ASSESS FOR SORTING PROCEEDURES
SIEVE INTO THREE FRACTIONS
EXTRA COARSE (>4.0 mm)
COARSE (1.0 - 4.0 mm)
FINE (500 µm - 1.0 mm)
IF MANY ANIMALS, SELECT SUB-SAMPLING PROCEDURE
PICK ENTIRE FRACTION UNDER SCOPE(10X)
PICK ENTIRE FRACTION UNDER SCOPE (10-20X)
PICK SUB-SAMPLES UNDER SCOPE (20X) record total # / sub-sample and place organisms into a common vial
CONTINUE SORTING SUBSAMPLES UNTIL TOTAL NUMBER > 300
PERFORM QA/QC PROCEEDURES ON EACH FRACTION
IDENTIFCATION OF ORGANISMS

5.22.3 Taxonomy and Reference Collections

The recommended level of identification for benthic invertebrates in freshwater is ***family*** for most aspects of the EEM program. For some phases, a lower taxonomic level may be recommended depending on the questions and objectives of the study (see Section 5.12). Lowest Practical Level (LPL) has been defined as genus for most insects and the lowest level possible without special procedures (dissection, microscopy) or reliance on specialists for other groups (Taylor 1997). Although this level of identification is not recommended for freshwater in general, the definition can be used as a guide if lower level identifications are desired for Focussed Monitoring or Investigation of Cause. Table 5.9 lists taxonomic references typically used for specific groups or regions.

Verification of taxonomy by recognized experts is also recommended for the reference collection. Museums are sometimes prepared to perform this service when remote areas are included in the study and new species or distribution records are likely. They will also sometimes serve as archives for selected samples or reference specimens. However, since considerable effort is involved in long-term maintenance of preserved biological material, the quantity of material submitted should be minimized.

Table 5-9: Taxonomic Keys for benthic invertebrate taxonomic identification in freshwater environments.

TAXON	Taxonomic Reference Typically Used
General Keys	Merritt and Cummins 1984, 1996; Peckarsky et al. 1990; Pennak 1978; Thorp and Covich 1991
Regional Keys	Clifford 1991 (Alberta) Lehmkuhl 1975a, 1975b, 1976, 1979 (Saskatchewan) Laplante et al. 1981 (Quebec)
Taxon Specific Keys	
ANNELIDA	
Oligochaeta	Brinkhurst 1986
Hirudinea	Klemm 1972, 1985, 1991.
Crustacea	Bousfield 1958
Amphipoda	Brandlova 1972
Decapoda	Dussart 1969
Cladocera	Crocker & Barr 1968
Copepoda	Fitzpatrick 1983
INSECTA	Chu and Cutkomp 1992; Hilsenhoff 1995
PLECOPTERA (stoneflies)	Fullington & Steward 1980; Harper & Hynes 1971; Hitchcock 1974; Stewart and Stark 1988
EPHEMEROPTERA (mayflies)	Allen & Edwards 1965; Bednarik & McCafferty 1970, 1979; Edmunds et al. 1976; Lewis 1974; Morihara and McCafferty 1979; McCafferty & Waltz 1990; Waltz 1994
ODONATA (dragonflies and damselflies	Hilsenfoff 1995; Westfall & May 1996; Walker 1933, 1955, 1958 Walker & Corbet 1978
TRICHOPTERA (caddisflies)	Schefter and Wiggins 1986; Schuster and Etnier 1978; Wiggins 1996
COLEOPTERA (beetles)	Brown 1992; Hilsenhoff & Schmude 1992
DIPTERA (flies)	Hilsenhoff 1995; Johannsen 1977; Oliver et al. 1978; Saether 1975, 1977; Simpson and Bode 1990; Wiederhom 1983, 1986 ; Wood 1963
GASTROPODA (snails)	Burch 1989; Clarke 1981
PELECYPODA (clams, mussels)	Mackie et al. 1980; Clarke 1981; Burch 1975 a,b; Mackie and Huggins 1983

Table 5.10 lists recommended levels of taxonomy desirable for major taxonomic groups of marine benthic organisms. In general, the level of taxonomy should be consistent in each major group for all samples from a survey, and from survey to survey. Organisms that cannot be identified to the desired level of taxonomic precision should be reported as a separate category in the fundamental dataset at the finest level of taxonomic resolution possible. Since the accuracy of the taxonomic work depends on the availability of up-to-date taxonomic literature, a basic library of identification keys is essential. Keys appropriate to the geographic region of study are recommended. A detailed list of taxonomic references for marine and estuarine habitat is found in Table 5.11. Microscope slide mounts should be prepared for taxa requiring detailed microscopic examination for identification. This may involve various steps including dissection, clearing, and staining. Slide preparation techniques are listed in Klemm et al. (1990). For marine benthos, juvenile or non-adult fauna should be identified and enumerated separately from adults, as they show different patterns of response to environmental effects. All identifications should be carried out or verified by a qualified and experienced taxonomist. Existing reference collection may be useful as well. An example is the Atlantic Reference Centre at Huntsman Marine Station in St. Andrews (N.B.). Photographic iconographs have been used to advantage (Camburn et al. 1984-1986).

Table 5-10: Recommended level of taxonomic precision for benthic invertebrates in marine environment

Taxon	Level
	class
Porifera	
Cnidaria	genus
Turbellaria	genus
Nemertea	genus
•Nematoda	phylum (not to be included in analyses)
Sipuncula	species
Priapulida	species
Brachiopoda	genus
Bryozoa	family
Mollusca	
• Aplacophora	genus
• Gastropoda	species
• Bia	species
• Polyplacophora	genus
• Scaphopoda	species
Annelida:	
• Polychaeta	species (except some immature)
• Oligochaeta	genus
Arthropoda:	
• Pycnogonida	family
• Cephalocarida	Sub-class
• Malacostraca	species
• •Copepoda	sub-class (remove from analyses)
• Cirripedia	species
•	
Ascidiacea	family
Echinodermata	species

•nematodes and harpacticoid copepods are meiofauna andonly a fraction of specimens will be captured by a 0.5 mm of 1 mm screen – therefore numbers are not representative and should be excluded from analyses (Holme and McIntyre 1984)

Table 5-11: List of marine and estuarine taxonomic benthic invertebrate keys for Canada

References
Abbot 1974 (seashells)
Abbott, R.T., et Sandström, G.F. 1982 (Mollusca)
Appy et al. 1980 (Bay of Fundy polychaetes)
Austin 1985 (Pacific invertebrates)
Baker 1980 (Tubificid species)
Banse 1972, 1974 (polychaetes)
Berkeley, 1952 a,b (Pacific Annelida)
Blake 1971 (Polydora, east coast)
Blake 1991 (Polychaeta, north Atlantic)
Blake, J.A. 1988 (Phyllodocidae (Polychaeta), Atlantic)
Bousfield 1960 (Atlantic seashells)
Bousfield and Hendryks 1994, 1995 a,b (Pacific Amphipoda)
Bousfield and Hoover 1995 (Pacific Amphipods)
Bousfield and Kendall 1994 (Pacific Amphipods)
Bousfield, E.L. 1973 (Amphipoda, Atlantic)
Brinkhurst 1982 (Oligochaetes)
Brinkhurst, R.O., and Baker, H.R. 1979 (Marine Tubificidae(Oligochaeta)
Brunel, P., Bossé, L. and Lamarche, G. 1998 (Catalog of the invertebrates of the Gulf of St.Lawrence)
Butler 1983 (Pacific shrimps)
Clark, H.L. 1907 (Holothuroidea)
Clark, H.L. 1915 (Ophiuroidea)
Coates 1980 (BC Enchytraeidae)
Coe, W.R. 1943(Nemertea, Atlantic)
Coe, W.R. 1972 (Echinodermata, Atlantic)
Cutler 1973 (Sipuncula)
Fauchald, K. 1977(Polychaeta)
Fournier and Petersen 1991 (Polychaeta)
Gibson, R. 1994 (Nemertea)
Gosner 1971
Graham, A. 1988 (Gastropoda)
Hart 1982 (BC crabs)
Hobson and Banse 1981 (BC polychaetes)
Hyman, L.H. 1941 (Polycladida (Turbellaria), Atlantic)
Hyman, L.H. 1944 (Turbellaria, Atlantic)
Keen and Coan 1974 (Mollusca)
Knight-Jones, P. 1978 (Spirorbidae (Polychaeta), Pacific and Atlantic)
Knight-Jones, P. 1983 (Sabellidae (Polychaeta)
Kozloff 1987 (Pacific N.W. Invertebrates)
Lambert 1981 (BC sea stars)
Laubitz 1972 (Caprellidae)
Light, W.J. 1977 (Spionidae (Polychaeta), Pacific)
Morris, P. A. 1975 (Mollusca, Atlantic)
Pettibone, M.H. 1963(Polychaeta, Atlantic)
Pettibone, M.H. 1992 (Pholoidea (Polychaeta)
Pettibone, M.H. 1993 (Polynoidae (Polychaeta)
Pohle, G.W. 1990 (Decapoda, Atlantic)
Sars, G.O. 1895 (Amphipoda)
Sars, G.O. 1899 (Isopoda)

Sars, G.O. 1900 (Cumacea)
SBMNH 1994a,b,c, 1995a,b,c, 1996a,b,c, 1997a,b
Schultz 1969 (isopod crustaceans)
Smith, R.I. 1964 (Marine invertebrate keys, Atlantic)
Squires, H.J. 1990 (Decapoda, Atlantic)
Steele and Brunel 1968 (Amphipoda)
Tattersall , W.M. 1951 (Mysidacea)
Thorp, J.H., and Covich, A.P . (Editors)1991(Freshwater invertebrate keys)
Ushakov, P.V. 1955 (Polychaeta)
Wallace 1919 (Bay of Funday Isopoda)
Watling, L. 1979(Cumacea, Atlantic)

5.22.4 QA/QC for Benthic Invertebrate Laboratory Operations

In the laboratory, invertebrate samples are processed and counts of the various taxa are made. It is recommended that the sorted, preserved samples from each survey be retained in an appropriate storage facility for at least 6 years, or until it is determined that no further information, will be required from the samples. It is important that all samples are processed in similar manners such that the amount of experimental error in the counts is minimized. To minimize processing error, the following items should be included in the QA/QC program:

1. all personnel involved in the sample processing and analyses should have appropriate training,
2. the effects of sub-sampling (if done) on abundance estimates should be examined on a minimum of 10% of the samples and the effects of sub-sampling on the sample estimates be documented,
3. re-sorting of randomly selected samples should be done to determine the success of initial sorting (see detailed discussions below)
4. appropriate taxonomic references should be used given the type of habitat and geographic location,
5. a complete reference collection for each mine should be compiled and verified by an external taxonomic expert, and updated when appropriate (i.e., when new taxa are recorded)
6. a system for archiving samples should be outlined,
7. detailed sample processing and laboratory notes should be maintained in a bound notebook.

Ecological sample processing involves, as a first step, sorting of organisms from debris and, possibly, sub-sampling sorted organisms for detailed identification. Inevitably, processing errors are associated with these activities and must be estimated (e.g. Kreis 1986, 1989). Verification of sorting efficiency is easily performed on a spot-check basis if the left-over debris from a sample is retained. It is recommended that at least 10 % of all samples be resorted, and that the criteria for an acceptable sort be that ≤ 10% of the total number of organisms were missed. If > 10% of the total number is found during the resort, then all the samples within that group of samples requires resorting. A further criterion which would require a resort is if an entire group of benthic invertebrates was missed by the sorter, (i.e. it wasn't recognized as an organism), even if the missed organism constituted < 10% of the total. The factors which should be considered when determining similar groups of samples include: 1) sampling area, 2) habitat class, and 3) individual sorters. The QA/QC guidelines apply independently to each group of samples sorted. Unsorted and sorted fractions are to be retained until taxonomy and sorting efficiency are confirmed.

For further information regarding all aspects of QA/QC procedures for benthic invertebrate programs refer to the 1999 AETE report (Beak 1999).

5.23 Tools for Alternative Monitoring Methods and/or Investigation of Cause

There are a number of possible tools for use as alternative monitoring approaches (Chapter 12) and tools for investigation of cause for the EEM program. The list of methods provided below is not meant to be exhaustive, and mines may propose additional scientifically defensible approaches. Tools should be cost-effective, recognised with primary literature documentation, readily available from consulting, academic or government laboratories, and applicable to the EEM program.

The text on alternative approaches and tools for the investigation of cause addresses the purpose, rationale, limitation and references for each method. A detailed description of the methods can be found in the references.

5.23.1 Use of Weight-of-Evidence Approaches to Establish Cause of Effect(s)

Distinguishing among the cumulative impacts of multiple stressors (which sometimes have confounding effects) requires establishing a definitive causality link to the mine effluent that is being evaluated. Environmental assessments of aquatic ecosystems are particularly prone to these impediments because these ecosystems often receive multiple, interacting effluent discharges. Assessments of monitoring results often rely, in large part, on field monitoring data that can only show correlation (rather than clear cause and effect) between the mine effluent and its presumed effect. Establishing a strong causal link, however, can benefit from a weight-of-evidence approach that combines information from a variety of sources. The strength of the causal link can be evaluated by using a formalized set of criteria originally developed in the field of epidemiology. These criteria are outlined in Table 5.12, which is based on Fox (1991), Suter (1993), Gilbertson (1997), Beyers (1998), Culp (1999), Culp et al. (2000), and Lowell et al. (2000). In many cases, not all of the criteria may be satisfied or the findings for some criteria may conflict with those for other criteria. In these cases, it can be useful to assign weights to each of the criteria in terms of their relative importance for a given assessment. In particular, assessments that include evidence from field and/or laboratory experiments will be stronger than those based on field monitoring alone. It should be noted that it often is not necessary to satisfy all of the criteria to provide a strong causal argument (e.g., see Beyers, 1998). The advantages to applying formalized criteria to assessments of the data generated by an EEM program include: (1) helping to tie together diverse assemblages of data on the effects of multiple stressors; (2) more clearly assigning causality; and (3) identifying important informational gaps (i.e., criteria that have not yet been addressed). This kind of approach can help to make assessments of mine effluent effects more rigorous and robust by combining alternative approaches and investigation of cause methods, resulting in more appropriate management actions.

Table 5-12: Formalized set of causal criteria forming part of a weight-of-evidence approach for assessment of mining effluent effects (Culp et al., 2000, Lowell et al., 2000).

Causal Criteria for Weight-of-Evidence Approach
Spatial correlation of stressor and effect along gradient from more to less exposure areas.
Temporal correlation of stressor and effect relative to time course of exposure.
Plausible mechanism linking stressor and effect.
Experimental verification of stressor effects under controlled conditions and concordance of experimental results with field data.
Strength: steep exposure/response curve.
Specificity: effect diagnostic of exposure to a particular stressor.
Evidence of exposure (contaminants or other indicators) in body of affected organisms.
Consistency of stressor/effect association among different studies within the region being studied.
Coherence with existing knowledge from other regions where the same or analogous stressors and effects have been studied.

References:

Beyers, D.W. 1998. Causal inference in environmental impact studies. J. N. Am. Benthol. Soc. 17:367-373.

Culp, J.M. 1999. Weight-of-Evidence Assessments for Large Rivers: Linking Community Bioassays with Field and Laboratory Studies. Workshop on Integrated Approaches for Interpreting Environmental Effects Monitoring Data. DOE EEM/1999/1.

Culp, J.M., R.B. Lowell and K.J. Cash. 2000. Integrating in situ community experiments with field studies to generate weight-of-evidence risk assessments for large rivers. Environ. Toxicol. Chem. 19(4):(*in press*).

Fox, G.A. 1991. Practical causal inference for ecoepidemiologists. J. Toxicol. Environ. Health 33:359-379.

Gilbertson, M. 1997. Advances in forensic toxicology for establishing causality between Great Lakes epizootics and specific persistent toxic chemicals. Environ. Toxicol. Chem. 16:1771-1778.

Lowell, R.B., J.M. Culp and M.G. Dube. 2000. A weight-of evidence approach for northern river risk assessment: integrating the effects of multiple stressors. Environ. Toxicol. Chem. 19(4):(*in press*).

Suter, G.W. II. 1993. Ecological Risk Assessment. Lewis Publishers, Boca Raton, FL, USA.

5.23.2 On-Site Community Bioassays Using Mesocosms as an Alternative Monitoring Method or Bioassessment Tool

Field-based community bioassays using mesocosms is an alternative approach to standard monitoring in order to determine if the effluent is having an effect on the benthic invertebrate community during Initial, Periodic, and Focussed Monitoring in situations of confounding environmental descriptors, multiple discharges, concern about safety to field crews, or habitat modifications from historical effluent discharges. This community bioassay approach can also be used for the Investigation of Cause when strong inference is necessary to make future decisions related to impact remediation. Experimentation using mesocosms provides replication and precise control over relevant variables, and provides a means for experimental manipulation of aquatic populations and communities under controlled exposure condition with the environmental relevance of ambient field conditions (i.e., temperature, water quality, photoperiod). This approach can offer a higher level of ecological complexity than acute toxicity tests, thereby providing important ecological information concerning the chronic effects of pollutants on food webs. Mesocosms range widely in design and several recent articles have reviewed their use and application in ecotoxicology (Shriner and Gregory 1984, Kosinski 1989 , Beyers an Odum 1993, Cooper and Barmuta 1993, Guckert 1993, Graney et al. 1994, Pontasch 1995, Culp et al. 1996). Recent examples of the use of mesocosms for effects assessment in Canada include Culp et al. (1996, 2000), Lowell et al. (1995), and Dube et al. (1998). Limitations to the application of mesocosms can include the lack of appropriate access and infrastructure requirements at test sites, the restriction to medium-term (i.e., weeks to months) testing such that longterm bioaccumulation processes cannot be examined, and (at the time of writing) a limited commercial availability of the technique. When applying field-based community bioassays using mesocosms as an alternative approach to standard bioassessments, it would be desirable to maintain the costs so as to be comparable to those of standard biological monitoring surveys.

See Section 12.1 for further details on the use of mesocosms.

References:

Beyers, R. J. and H. T. Odum. 1993. *Ecological microcosms.* Springer-Verlag, New York, NY, USA.

Cooper, S. D., and L. Barmuta. 1993. Field experiments in biomonitoring. In D. M. Rosenberg and V. H. Resh, eds., *Biomonitoring and benthic invertebrates.* Chapman and Hall, New York, NY, USA, pp. 399-441.

Culp, J.M. 1999. Weight-of-Evidence Assessments for Large Rivers: Linking Community Bioassays with Field and Laboratory Studies. Workshop on

Integrated Approaches for Interpreting Environmental Effects Monitoring Data. DOE EEM/1999/1.

Culp, J. M. and C. L. Podemski. 1996. Design and application of a novel stream microcosm system for assessing effluent impacts to large rivers. In M. R. Servos, K. R. Munkittrick, J. H. Carey and G. J. Van Der Kraak, eds., *Environmental fate and effects of pulp and paper mill effluents.*, St. Lucie Press, Boca Raton, FL, USA, pp. 549-555.

Culp, J. M., C. L. Podemski, K. J. Cash, and R. B. Lowell. 1996. Utility of field-based artificial streams for assessing effluent effects on riverine ecosystems. *J. Aquat. Ecosystem Health* 5:117-124.

Culp, J.M., R.B. Lowell and K.J. Cash. Integrating in situ community experiments with field studies to generate weight-of-evidence risk assessments for large rivers. Environ. Toxicol. Chem. 19(4):(*in press*).

Dube, M. G. and J. M. Culp. 1996. Growth responses of periphyton and chironomids exposed to biologically treated bleached-kraft pulp mill effluent. *Environ. Toxicol. Chem.* 7:2019-2027.

Dube, M.G., J.M. Culp and J.A. Smith. 1998. Development and testing of a mesocosm technology for environmental effects monitoring (EEM) methods application. pp. 43-46. *In* S.C. Courtenaay, W.R. Parker and G.P. Rawn [eds.] Proceedings of a a workshop to assess alternatives to the fish survey component of the environmental effects monitoring program for Canadian pulp and paper mills. Can. Tech. Rep. Fish. Aquat. Sci. No. 2233.

Graney, F.L., J.J. Kennedy, and J.H. Rodgers. (eds.) 1994. Aquatic mesocosm studies in ecological risk assessment. Special publication of SETAC. Press, Pensacola, FL, USA.

Guckert, J. B. 1993. Artificial streams in ecotoxicology. In G. A. Lamberti and A. D. Steinman, eds., *Research in artificial streams: applications, uses, and abuses, J. North Amer. Benthological Soc.* 12:313-384, pp. 350-356.

Kosinski, R. J. 1989. Artificial streams in ecotoxicological research. In A. Boudou and F. Ribeyre, eds., *Aquatic ecotoxicology: fundamental concepts and methodologies*, Vol. I, CRC Press, Boca Raton, FL, USA, pp. 297-316.

Lowell, R. B, J. M. Culp and F. J. Wrona. 1995. Stimulation of increased short-term growth and development of mayflies by pulp mill effluent. *Environ. Toxicol. Chem.* 14:1529-1541.

Podemski, C. L. and J. M. Culp. 1996. Nutrient and contaminant effects of bleached kraft mill effluent on benthic algae and insects of the Athabasca River, AB, Canada. In M. R. Servos, K. R. Munkittrick, J. H. Carey and G. J. Van Der Kraak, eds., *Environmental fate and effects of pulp and paper mill effluents,* St. Lucie Press, Boca Raton, FL, USA, pp. 571-580.

Pontasch, K. W. 1995. The use of stream microcosms in multispecies testing. In J. Cairns and B. R. Niederlehner, eds, *Ecological toxicity testing: scale, complexity and relevance*, Lewis Publishers, Boca Raton, FL, USA, pp. 169-191.

Rodgers, J. H., Jr., N. O. Crossland, E. R. Kline, W. B. Gillespie, Jr., R. A. Figueroa, and P. B. Dorn. 1996. Design and construction of model stream ecosystems. *Ecotoxicol. Environ. Safety* 33:30-37.

Shriner, C. and T. Gregory. 1984. Use of artificial streams for toxicological research. *Crit. Rev. Toxicol.* 13:353-281.

5.23.3 Methods for Determining Biomarkers and Metal Body Burdens

The use of biomarkers and metal body burdens can be applied during Focussed Monitoring (to determine magnitude) and Investigation of Cause when an effect has been identified but previous monitoring has failed to provide a satisfactory explanation of cause. Significant increases in metal body burdens or in biomarker responses (either over time or relative to reference sites) will indicate that metals are bioavailable in the receiving environment. If data on the relationship between bioaccumulation and toxicity are available, metal bioaccumulation can be used to directly infer the cause of toxicity. In the domain of biomarkers for metals, metallothionein (MT) conforms to most of the criteria defined for such tools. For example, MT responds specifically in a dose-dependent manner to changes in ambient levels of a trace metal or of a group of trace metals (Cd, Cu, Zn, Ag). In addition, high levels of MT are associated with deleterious effects on organisms and populations because these high concentrations are indicative of metal detoxification failure and of non-specific metal binding to cellular targets of toxicity. Appropriate monitoring species for MT measurements and metal body burdens include freshwater and marine bivalves, benthic insect larvae and amphipods. Indigenous specimens can be collected at exposure sites, or they can be transplanted from reference locations to these exposure sites. Bioaccumulation is particularly useful for identification of bioavailability of non-essential metals (Cd, Pb, Tl, etc), but copper and zinc in invertebrate tissues may be regulated to varying degrees, and their background concentrations may be high. This may limit the usefulness of copper and zinc measurements in invertebrates. There is yet no field evidence that MT concentrations respond to metals like As, Ni, Pb, and Hg. In addition, the MT biomarker will not track exposure to non-metallic chemicals such as ammonia and cyanide. Evidence is accumuiating in the literature to the effect that MT constitutes an excellent biomarker of metal-induced effects. However, the effect-biomarker approach is still in the research mode (at the time of writing) and awaits a field validation/demonstration in an independent set of mining area sites. Body metal and MT concentrations are easy and relatively inexpensive to measure, and these analyses are available in the private sector (body metal more than MT).

References:

Metal body burdens

Borgmann, U. and Norwood, W.P. 1997. Toxicity and accumulation of zinc and copper in *Hyalella azteca* exposed to metal-spiked sediments. Can. J. Fish. Aquat. Sci. 54: 1046-1054.

Borgmann, U. and Norwood, W.P. 1997. Identification of the toxic agent in metal-contaminated sediments from Manitouwadge Lake, Ontario, using toxicity-accumulation relationships in *Hyalella azteca*. Can. J. Fish. Aquat. Sci. 54: 1055-1063.

Borgmann, U.2000. Methods for assessing the toxicological significance of metals in aquatic ecosystems: Bioaccumulation-toxicity relationships, water concentrations, and sediment spiking approaches. Aquat. Ecosyst. Health Management. In press.

Borgmann, U. and W.P. Norwood. 1999. Predicting the toxicity of lead in sediments to *Hyalella azteca*: the significance of bioaccumulation and dissolved metal. Can. J. Fish. Aquat. Sci. 56: 1494-1503.

Tessier, A., Couillard, Y., P.G.C. Campbell, and J.C.A. Auclair. 1993. Modeling Cd partitioning in oxic lake sediments and Cd concentrations in the freshwater bivalve *Anodonta grandis*. Limnol. Oceanogr. 38: 1-17.

Metallothionein

Couillard Y. 1997. With the collaboration of L. St-Cyr. Technical evaluation of metallothionein as a biomarker for the mining industry. Technical report for the Aquatic Effects Technology Evaluation (AETE) program, AETE Project 2.2.1, prepared for Natural Resources Canada, Canada Centre for Mineral and Energy Technology (CANMET). 191 p.

Couillard, Y., A. Giguère, P.G.C. Campbell, O. Perceval,B. Pinel-Alloul, and L. Hare. 1999. Field evaluation of the use of metallothionein as a biomarker for metal contamination and toxic effects in the freshwater bivalve *Pyganodon grandis*: Subcellular metal partitioning. SETAC 20th Annual Meeting, 14-18 November 1999, Philadelphia. PA.

Giguère, A., Y. Couillard, O. Perceval, P.G.C. Campbell, J. Pellerin-Massicotte, and B. Pinel-Alloul.1999. Field evaluation of the use of metallothionein as a biomarker for metal contamination and toxic effects in the freshwater bivalve *Pyganodon grandis*: Responses at the organism level. SETAC 20th Annual Meeting, 14-18 November 1999, Philadelphia. PA.

Gillis, P.L., D.G. Dixon, T.B. Reynoldson, and L.C. Diener. 1999. Metallothionein as a biomarker for trace metal exposure and effects in *Tubifex tubifex* and *Chironomus riparius*.
SETAC 20th Annual Meeting, 14-18 November 1999, Philadelphia. PA.

Laflamme J.S., Y. Couillard, P.G.C. Campbell and A. Hontela. 1998. Impaired physiological response to stress and high levels of metallothionein in perch (*Perca flavescens*) from lakes contaminated by heavy metals. 25th Annual Aquatic Toxicity Workshop, 18- 21 October 1998, Quebec City, QC.

Perceval, O., Y. Couillard, A. Giguere, B. Pinel-Alloul, and P.G.C. Campbell. 1999. Field evaluation of the use of metallothionein as a biomarker for metal contamination and toxic effects in the freshwater bivalve *Pyganodon grandis*: Responses at the population level. SETAC 20th Annual Meeting, 14-18 November 1999, Philadelphia. PA.

Pinel-Alloul, B., P.G.C. Campbell, L. Hare, Y. Couillard, A. Giguere, and O. Perceval. 1999. Field evaluation of the use of metallothionein as a biomarker for metal contamination and toxic effects in the freshwater bivalve *Pyganodon grandis*: Overview. SETAC 20th Annual Meeting, 14-18 November 1999, Philadelphia. PA.

Wang, D., Y. Couillard, P.G.C. Campbell and P. Jolicoeur. 1999. Changes in subcellular metal partitioning in the gills of freshwater bivalves (*Pyganodon grandis*) living along an environmental cadmium gradient. Can. J. Fish. Aquat. Sci. 56: 774-784.

5.23.4 Lethal and Sublethal Toxicity Tests

Lethal and sublethal toxicity test methods can be applied during Focussed Monitoring and Investigation of Cause when an effect has been identified, or when previous work failed to provide a satisfactory explanation of cause. These methods provide a direct determination of lethal or sublethal toxicity and can verify that alterations in benthos are due to toxicity of the mine effluent rather than confounding factors. For example, adverse effects on benthic community structure may be due to factors other than effluent toxicity, including differences in environmental regime. Concurrent impairment of benthic community structure and toxicity implicates the effluent itself as the cause of changes in the benthos. These methods also provide important information for interpretation of field effects in situations where benthic community data are inconclusive, or if only pollution tolerant species are present in both impacted and reference sites. For example, because toxicity of metals in sediments is not proportional to total metal concentrations, bioavailability can vary greatly from one sediment to another. Thus, metal concentrations in invertebrates exposed to sediments can be a direct measure of bioavailability, and if data on the relationship between bioaccumulation and toxicity are available, metal bioaccumulation can be used to directly infer the cause of toxicity. Alternatively, metal concentrations in overlying water during static toxicity tests with sufficiently large water to sediment ratios can also be used to estimate metal bioavailability and the cause of sediment toxicity. Note that bioaccumulation is particularly useful for identification of bioavailability of non-essential or non-regulated metals (Cd, Tl, etc.). However, bioaccumulation of copper and zinc in invertebrate tissues is less useful because these elements are regulated to varying degrees, and their background concentrations can be high. For regulated metals the overlying water is sometimes a more useful indicator of bioavailability and cause of toxicity.

There is an array of laboratory test methods available, some of which are listed below. In addition, the use of whole sediment toxicity, sediment contaminant analysis, and benthic community survey assessment at a common sediment depositional location allows for data correlations between the different sediment monitoring tools through the application of the Sediment Quality Triad. Strong correlations provide an indication of the likelihood of sediment contaminants being involved in any observed field effects. In addition to laboratory tests of media collected in the field, tests can also be conducted in the field with caged animals. Field transplants of invertebrates are complementary approaches to the biomonitoring of benthic communities, as these methods will help interpret results in situations of confounding factors, multiple discharges, or habitat modifications from historical effluent discharges. These methods may also be useful in identifying the cause for the absence of taxa at exposure sites.

Lethal and sublethal toxicity tests have a number of limitations that must be considered. Because laboratory toxicity test methods do not determine toxicity in situ, they might not account for toxicity due to overlying water, or for the effects of overlying water chemistry on contaminant toxicity. These problems will be most severe when the chemistry of the overlying water used in the toxicity test differs substantially from in situ water, or when conditions in the field or during toxicity tests are such that contaminant concentrations in overlying water are far from equilibrium with the sediment. For field transplant approaches, experimental sites should be visited at least twice and sites must be protected from vandalism. Cage bioassays are often limited to the use of older (larger) and less sensitive individuals because cage mesh size must be large enough to allow adequate water exchange to prevent cage effects such as fouling, sedimentation and low dissolved oxygen concentration. When applying laboratory or sediment toxicity tests and bioassays, it would be desirable to maintain the costs of these methods so as to be comparable to the costs of a standard survey at the site. In addition, the methods should be readily available from consulting, academic or government laboratories.

References:

Methods: whole sediment survival and growth or reproduction tests with Hyalella azteca Chironomus sp., Tubifex tubifex, and Hexagenia sp.; measurement of body concentrations of metals in invertebrates used in sediment toxicity tests

AETE (Aquatic Effects Technology Evaluation Program). 1996. 1995 Field evaluation of aquatic effects monitoring methods - Pilot study, Volume 1, AETE Project 4.1.1. Canada Centre for Mineral and Energy Technology, Ottawa, ON.

AETE (Aquatic Effects Technology Evaluation Program). 1998. Summary and cost-effectiveness of aquatic effects monitoring technologies applied in the 1997 AETE field evaluation program. Canada Centre for Mineral and Energy Technology, Ottawa, ON.

ASTM (American Society for Testing and Materials). 1997. Standard test methods for measuring the toxicity of sediment-associated contaminants with fresh water invertebrates (Annex 3, *Hexagenia* spp., and Annex 4, *Tubifex tubifex*). Volume 11.05, Standard E1706-95b.

Bedard, D., Hayton, A., and Persaud, D. 1992. Ontario Ministry of Environment laboratory sediment biological testing protocol. Water Resources Branch, Ontario Ministry of Environment, Toronto, ON, 26 pp.

Borgmann, U. and Norwood, W.P. 1997. Toxicity and accumulation of zinc and copper in *Hyalella azteca* exposed to metal-spiked sediments. Can. J. Fish. Aquat. Sci. **54**: 1046-1054.

Borgmann, U. and Norwood, W.P. 1997. Identification of the toxic agent in metal-contaminated sediments from Manitouwadge Lake, Ontario, using toxicity-accumulation relationships in *Hyalella azteca*. Can. J. Fish. Aquat. Sci. 54: 1055-1063.

Borgmann, U. 2000. Methods for assessing the toxicological significance of metals in aquatic ecosystems: Bioaccumulation-toxicity relationships, water concentrations, and sediment spiking approaches. Aquat. Ecosyst. Health Management. In press.

Borgmann, U. and W.P. Norwood. 1999. Predicting the toxicity of lead in sediments to *Hyalella azteca*: the significance of bioaccumulation and dissolved metal. Can. J. Fish. Aquat. Sci. 56: 1494-1503.

Borgmann, U. and W.P. Norwood. 1999. Sediment toxicity testing using large water-sediment ratios: an alternative to water renewal. Environ. Poll. 106: 333-339.

Borgmann, U., V. Cheam, W.P. Norwood and J. Lechner. 1998. Toxicity and bioaccumulation of thallium in *Hyalella azteca*, with comparison to other metals and prediction of environmental impact. Environ. Poll. 99: 105-114.

Borgmann, U., R. Néron and W.P. Norwood. 2000. Quantification of bioavailable nickel in sediments and toxic thresholds to Hyalella azteca. Environ. Poll. submitted.

Deaver, E. and Rodgers, J.H., Jr. 1996. Measuring bioavailable copper using anodic stripping voltammetry. Environ. Toxicol. Chem. 15: 1925-1996.

EC (Environment Canada). 1997. Test for survival and growth in sediment using the larvae of freshwater midges (*Chironomus tentans* or *Chironomus riparius*). Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/32.

EC (Environment Canada). 1997. Test for survival and growth in sediment using the freshwater amphipod *Hyalella azteca*. Environmental Protection Service, Ottawa, ON, Report EPS 1/RM/33.

MacLean, R.S., Borgmann, U., and Dixon, D.G. 1996. Bioaccumulation kinetics and toxicity of lead in *Hyalella azteca* (Crustacea, Amphipoda). Can. J. Fish. Aquat. Sci. 53: 2210-2220.

Neumann, P.T.M., Borgmann, U., and Norwood, W.P. 1999. The effect of gut clearance on metal body concentrations in *Hyalella azteca*. Environ. Toxicol. Chem. 18: 976-984.

Warren, L., Tessier, A, and Hare, L. 1998. Modeling cadmium accumulation by benthic invertebrates in situ: the relative contributions of sediment and overlying water compartments to organism cadmium concentrations. Limnol. Oceanogr. 43: 1442-1454.

Methods: Field bioassays using caged animals

Couillard, Y., P.G.C. Campbell, A. Tessier, J. Pellerin-Massicotte, and J.C. Auclair. 1995. Field transplantation of a freshwater bivalve, *Pyganodon grandis*, across a metal contamination gradient. I. Temporal changes in metallothionein and metal (Cd, Cu, Zn) concentrations in soft tissues. Can. J. Fish. Aquat. Sci. 52: 690-702.

Couillard, Y., P.G.C. Campbell, J. Pellerin-Massicotte, and J.C. Auclair. 1995. Field transplantation of a freshwater bivalve, *Pyganodon grandis*, across a metal contamination gradient. II. Metallothionein response to Cd and Zn exposure, evidence for cytotoxicity, and links to effects at higher levels of biological organization. Can. J. Fish. Aquat. Sci. 52: 703-715.

Henry, C.J., Higgins, K.F., and Buhl, K.J. 1994. Acute toxicity and hazard assessment of Rodeo®, X-77 Spreader®, and Chem-Trol® to aquatic invertebrates. Arch. Environ. Contam. Toxicol. 27: 392-399.

Muir, D.C.G., Kenny, D.F., Grift, N.P., Robinson, R.D., Titman, R.D., and Murkin, H.R. 1991. Fate and acute toxicity of bromoxynil esters in an experimental prairie wetland. Environ. Toxicol. Chem. 10: 395-406.

Mulliss, R.M., Revitt, D.M., and Shutes, R.B.E. 1996. A statistical approach for the assessment of the toxic influences on *Gammarus pulex* (Amphipoda) and *Asellus aquaticus* (Isopoda) exposed to urban discharges. Wat. Res. 30: 1237-1243.

Stewart, R. 1997. Technical evaluation of molluscs as a biomonitoring tool for the canadian mining industry. AETE Project 2.3.1. Aquatic Effects Technology Evaluation (AETE) Program, Canada Centre for Mineral and Energy Technology, Ottawa, ON. 248 p.

van Wijngaarden, R.P.A., van den Brink, P.J., Crum, S.J.H., Voshaar, J.H.O., Brock, T.C.M., and Leeuwangh, P. 1996. Effects of the insecticide Dursban® 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: I. Comparison of short-term toxicity between the laboratory and the field. Environ. Toxicol. Chem. 15: 1133-1142.

5.23.5 Analysis of Sediment Cores for Historical Trends

Sedimentary records from depositional areas of waterbodies can be used to indicate limnological conditions in recent and ancient history (Frey 1998). Precise dating of sediments, combined with an inventory of the remains of certain organisms and plant material (e.g. diatoms, zooplankton, insects), provide a chronology of changes that often can be linked to the period of anthropogenic influence. In addition to the waterbody itself, the history of the watershed and airshed may be deduced, and the influences of natural events may be distinguished from anthropogenic impacts. A substantial volume of literature is available on the subject, with a useful synthesis of the science provided by Frey (1998). Due to the level of expertise needed to undertake this type of analysis, the availability of paleolimnological services is limited. In addition, the analyses are restricted to resolving trends over longer time frames (multiple years to decades) as a result of sedimentation processes such as bioturbation. Costs of the technique will be site-specific.

Frey, D. G. 1998. What is paleolimnology? J. Paleolimnology 1:5-8.

5.24 Other Benthic Invertebrate Measures and Organisms

Benthic invertebrates are recommended as the primary indicator organisms for use in an EEM program for monitoring effects on fish habitat. However, the level of identification and measures recommended in the main text of the guidance document are not an exclusive list of measures for which benthic invertebrates can be evaluated. Additional measures include biomass, lower level identification, secondary production, and population fitness parameters. Several recent reports outline the usefulness of some of these additional parameters for monitoring programs (Taylor 1997; Taylor and Bailey 1997; Feltmate and Fraser 1998).

Benthic invertebrate biomass in **marine environments** can provide additional useful information because it is related to the availability of energy to other trophic levels (e.g., fish). For marine communities, some investigators suggest that an analysis of benthic abundance and biomass together provide a sensitive indicator of changes in the composition of the benthic community (e.g., Warwick 1986; Warwick et al. 1987; Clarke 1990; Burd et al. 1990). For example, in marine samples, it is in the measurement of distributions of biomass that the three main functional groups of benthic organisms, microfauna (grain surface dwellers), meiofauna (interstitial organisms) and macrofauna (burrowers and epifauna) can be distinctly separated (Schwinghamer 1981, 1983). Because these three groups of organisms have different reproductive modes, metabolic rates, life-histories and habitat adaptations, they respond differently to habitat perturbation. This could be particularly important in arctic subtidal habitats, where abundance may be low, but individuals may be large (ref.)However, because precise biomass measurements are time consuming and problematic (c.f. Crisp 1984) unless collected in more detail and more often than is feasible for EEM requirements, it is only possible to determine relative changes in biomass of samples for the EEM surveys. This can be simply done by taking blotted wet-weight measurements of representative sized adult specimens of each species for each survey. Since the method is non-destructive, the reference collection may be used for this purpose prior to external verification or archiving. The mean weight of a given species can then be used to transform species abundance data to relative species biomass data for further summary or statistical analyses. These data show relative, large scale changes only, and cannot be used to infer production or trophic flow rates within benthic communities.

In addition to benthic invertebrates, several other types of aquatic biota were also considered for use in the EEM program. The most relevant ones were 1) phytoplankton, 2) macrophytes, and 3) periphyton. Each of these was evaluated on the basis of:

1) Sensitivity
    a) ability to distinguish exposure and reference areas
    b) early warning
2) Ecological relevance
3) Validity and repeatibility
    a) peer review and validation in laboratory and field
    b) temporal variability
4) Site specificity and applicability
    a) applicability in a wide range of environmental conditions and geographic locations
    b) ability to differentiate between point and non-point sources

5) Costs
- a) average monetary cost
- b) commercial availability

Based on these criteria, these other biotic categories were deemed inappropriate for use as a primary indicator. Each of them, however, shows promise for use in some situations and may, therefore, be considered for inclusion in specific monitoring programs as an additional tool or indicator on a site-by-site basis.

The following screening evaluations are largely based on findings for freshwater habitats from the AETE program and highlight the strengths and weaknesses of each group. The use of macrophytes for marine environment is only feasible in intertidal, hard substrate areas. If this is deemed to be a major habitat which is potentially affected by mining effluents, then shoreline quadrat studies of percent coverage and biomass per unit area, major species, and possibly associated epifauna such as amphipods may be appropriate as supportive indicators (see Cross 1981).

SCREENING EVALUATION

TEST: PHYTOPLANKTON - SPECIES COMPOSITION, DENSITY AND BIOMASS

Background : Structure of phytoplankton communities can be altered due to environmental stressors that affect certain sensitive species and not others. This approach has been used widely to examine the affects of metal pollution and has been applied to long term monitoring progress at mine sites. Standing crop and biomass (usually measured as chlorophyll a) of phytoplankton communities are typically assessed in any study in which primary producers are considered and have been used as indicators in long-term environmental programs at mine sites (Golan 1996)

1.Sensitivity	a) Can the indicator detect differences between impact and reference areas? ▪ Multivariate analyses have successfully illustrated effects of metals on phytoplankton communities but data from large numbers of lakes or sampling points are required. Variability in diversity index values make them less useful than simple species richness numbers. ▪ On a long term basis, biomass does not appear to be a useful measure of the effects of metal contamination on primary producers since long term shifts in species composition or reduction in grazing pressure can counteract any short term depression in algal biomass (Yan 1979). Additionally, dispersal of phytoplankton between reference and impact areas is a potential problem and often there is no relationship between contaminant levels and biomass. b) Will the indicator provide an early warning of possible changes between impact and reference areas? ▪ Schlinder (1987) suggests that measuring species changes in phytoplankton may serve as an early indicator of environmental stress. As several phytoplankton species are tolerant or intolerant of metals, presence /absence data may serve as an early warning to metal pollution but must be compared with a reference data set. Consistent patterns of tolerance or intolerance to metal pollution are found in some species (lists provided in Table 2.1, AETE report 2.3.2), and this approach has been used successfully at a number of metal contamination sites (Yan 1979, Austin and Munteanu 1984, Oliverina 1985) ▪ Although species composition may be an early warning indicator, biomass or density changes may only be apparent immediately after the input commences.
2.Ecological Relevance	

3.Validity and Repeatability	a) Has the indicator been peer reviewed and validated in the laboratory and the field? ■ Effects on the composition of natural phytoplankton communities by metals have been demonstrated in long term *in situ* experiments (Gachter and Nares 1979, Denoyelles et al. 1980). Results from these experiments concur with field studies where multivariate analyses have been used to infer relationships (Dixit et al. 1989, 1991; Nicholls et al. 1992) between contaminants and species composition. ■ Although short term experiments show metal enrichment will reduce standing crop and photosynthesis of phytoplankton (Blinn et al. 1977, Cote 1983, Takamura et al. 1989, Havens 1994), longer term studies indicate that density and biomass tend to be similar between impact and reference areas (Thomas and Seibert 1977, Gachter and Mares 1979, Autin and Munteanu 1984) **b)** Does the indicator vary seasonally or yearly? ■ Phytoplankton communities experience seasonal patterns in species composition, density and biomass. Therefore, species absence and standing crop data must be viewed with caution.
4.Site Specificity and Applicability	a) Can the indicator be used in a wide range of conditions at geographically different locations? ■ It is difficult to generalize from site to site regarding the effects of metals on phytoplankton communities. Many studies show similar effects but the particular species affected are typically site dependent. There are some species which are consistently tolerant (e.g., *Rhizosolenia*) over a broad geographical distribution (Deniseger et al. 1990, Bonacina et al. 1992, Havas et al. 1995), but others exhibit site-specific responses (Foster 1982, Whilton 1984). Lab studies indicate uptake of metals by algae is metal and pH dependent, i.e., algal metal accumulation is not readily related to external concentrations (Cain et al. 1980). So, although phytoplankton are abundant and widespread, local variability can be substantial. To use phytoplankton as a biomonitoring tool, it is essential to have appropriate reference areas. **b)** Can the indicator differentiate between localized impacts (point source) and generalized (non-point source) impacts? ■ Although gradients in phytoplankton species composition have been detected across several bodies of water, the potential problems with dispersal of phytoplankton between reference and exposure areas would likely be a substantial problem for either point or non-point source impacts.

5.Costs	a) What is the average cost of using the method (including QA/QC considerations)? ▪ $20-30 per sample for chlorophyll a ▪ $80-150 per sample for generic id (cell counts only) **b)** Is the indicator readily and/or commercially available? ▪ There are a limited number of experienced algal taxonomists available and turn around time must be considered. The analysis of chlorophyll a as a measure of density and biomass can be more easily performed in many laboratories with the appropriate equipment.

SCREENING EVALUATION

TEST: MACROPHYTES- SPECIES AND COMMUNITY COMPOSITION

Background : Streams and lakes exposed to high levels of metals can show differences in the standing crop, diversity and presence/absence of rooted plant communities when compared to reference areas (St-Cyr et. al. 1997).

1.Sensitivity	a) Can the indicator detect differences between impact and reference areas? ▪ The absence of macrophytes from impact areas where they would normally be expected (i.e., in reference areas) can be regarded as an indication of pollution (Gorham and Gordon 1963, Besch and Roberts-Pichette 1990, Yan et al. 1985). Measures of species abundance, diversity and richness can reflect general toxicity effects of potential contaminants in the system (Hellawell 1986). In fact, Gorham and Gordon (1963) have observed reductions in species diversity of submerged and floating macrophytes with decreasing distance from sulfur dioxide smelters in Sudbury. Interestingly, since emissions were reduced, recovery has occurred with increases in species richness observed (Ripley et al. 1996). In other studies, downstream recovery of macrophyte flora has been observed below lignite mines (Sand-Jensen and Rasmussen 1978) although results can be confounded by natural factors such as substrate, depth, slope, etc., b) Will the indicator provide an early warning of possible changes between impact and reference areas? ▪ Macrophyte community composition appears to be more of a long-term indicator of environmental stress rather than a tool for early detection of potential impacts.
2.Ecological Relevance	a) Can the indicator show that differences between impact and reference areas are ecologically relevant to the ecosystem or one of its components. ▪ Macrophytes are sedentary and provide a measure of the impacts of the bioavailable fraction of metals over a time integrated picture (Whitton et al. 1981). Impacts on macrophyte density and biomass could substantially affect the other members of the freshwater community as they may provide habitat for higher trophic levels within the community.

3.Validity and Repeatability	a) Has the indicator been peer reviewed and validated in the laboratory and the field? ▪ Many studies have demonstrated a reduction in species diversity and abundance in macrophytes exposed to metal-polluted waters (Gorham and Gordon 1963, Besch and Robert-Pichette 1970, Kelly 1988, Ribpley et al. 1986). Kelly (1988) has compiled a list of species along a contaminant gradient and found changes in composition, decreases in richness closer to the source. Further, Haslam (1982) and Small et al. (1996) have developed an elaborate method of monitoring macrophytes along a pollution gradient. Although a few attempts have been made to develop routine methods for practical applications (Kelly 1988; Sortkjaer 1984), further work needs to be done before the routine use of macrophytes as biomonitor organisms can be effectively applied in the field **b)** Does the indicator vary seasonally or yearly? ▪ Macrophyte populations vary seasonally in most rivers and lakes with marked seasonal variations in biomass. In fact, macrophytes are really only available as a bio-monitor during a relatively short growing season (Hellawell 1986).
4.Site Specificity and Applicability	a) Can the indicator be used in a wide range of conditions at geographically different locations? ▪ Whitton et al. (1981) suggested a list of ten macrophyte species to be used as standard bio-indicators. However, for Canadian studies, there are only 3 frequently encountered species recommended as potential biomonitor candidates in Canadian environments. **b)** Can the indicator differentiate between localized impacts (point source) and generalized (non-point source) impacts? ▪ no information or unknown?

.	a) What is the average cost of using the method (including QA/QC considerations)? ▪ no information or unknown? b) the indicator readily and/or commercially available? ▪ Macrophyte field sampling often is difficult because S.C.U.B.A. diving may be required. Laborartory analysis is simple since macrophytes are easy to identify and quantify and do not require extensive expertise. .

SCREENING EVALUATION

TEST: PERIPHYTON – SPECIES COMPOSITION, DENSITY AND BIOMASS

Background : Structural measurements of periphyton generally involve counts of organisms, providing information on abundance, species richness and community structure, whereas functional measures study changes in primary production, respiration and detrital processing (Clements 1991).

1.Sensitivity	a) Can the indicator detect differences between impact and reference areas? ▪ Assessments of community structure in monitoring programs have clearly demonstrated that the species composition at metal impacted sites can be different than that at reference sites (Clements 1991). However, confounding factors and natural changes in stream characteristics can make comparisons between reference and impact areas very difficult and/or complicated. For example, the response of diatoms to metals can be inconsistent, with some species being classified as resistant in one location (Yan 1979) and sensitive in another (Oliveria 1985, Bonacina et al. 1988). b) Will the indicator provide an early warning of possible changes between impact and reference areas? ▪ Organisms within the periphyton mat have relatively high turn over rates. Hence, they are among the first organisms to respond to environmental stress, and among the first to recover. Sensitive species are often replaced by more tolerant species (Austin 1983, Austin et al. 1985). As a result, periphyton communities often reflect the current environmental conditions (Lowe and Pan 1996) and could be used as an early warning indicator.

2.Ecological Relevance	a) Can the indicator show that differences between impact and reference areas are ecologically relevant to the ecosystem or one of its components. ▪ Periphyton is generally non-mobile, and integrates effects of environmental variables in a relatively short time frame. The microalgal component of periphyton is important as both a primary producer in rivers and littoral zones in lakes (Cattaneo and Kalff 1980) and as a food source for invertebrates(Whitton 1984). As a food source for higher trophic levels, periphyton potentially plays an important role in contaminant transfer between trophic levels. Also due to its sedentary nature, periphyton is a good indicator of local conditions and plays a key role at the interface between substratum and surrounding water, perhaps by influencing biogeochemical pathways and dynamics.
3.Validity and Repeatability	a) Has the indicator been peer reviewed and validated in the laboratory and the field? ▪ Periphyton has been examined for water quality assessments in different systems by a large number of investigators (Chessman 1985, Clements and Kiffney 1994, Cattaneo et al. 1995). Various sampling methods have been used, including scraping, brushing or aspiration. Periphyton communities are often species-rich and spatially compact relative to other aquatic groups, and representative samples can be collected from a few square centimeters of substratum (Lowe and Pan 1996). However, some authors have suggested that there is difficulty obtaining representative and uniform samples because of spatial heterogeneity and difficulty sampling organisms (Wietzel et al. 1979). Regardless of potential sampling problems, many studies have characterized algal communities in streams and have related abundance of specific taxa to the presence of metals (Clements 1991; Whitton 1984). Studies in lakes have similarly identified shifts in species composition with loss of metal sensitive species (Austin and Deniseger 1985, Austin et al. 1985, Roch et al. 1985). **b)** Does the validity of the indicator vary seasonally or yearly? ▪ Different responses of species to metal contamination may be due to seasonal differences and while temporal comparisons are rarely possible, spatial comparisons are of considerable value (Stokes 1983)

4.Site Specificity and Applicability	a) Can the indicator be used in a wide range of conditions at geographically different locations? ▪ Periphyton biomass, density and species composition are strongly influenced by local conditions, and some authors (Kutka and Richards 1986; St-Cyr 1997) have concluded that the spatial heterogeneity and lack of methodological standardization may restrict its use in impact studies of mine effluents. However, controlled exposures in experiments suggest periphyton may be an effective indicator of metal impacts Furthermore, Lowe and Pan (1996) concluded that benthic algal monitoring data would be most valuable when combined with a suite of additional monitoring data including physical and chemical measurements and the analysis of other biota of aquatic communities such as invertebrates. **b)** Can the indicator differentiate between localized impacts (point source) and generalized (non-point source) impacts? ▪ no information or unknown?
5.Costs	a) What is the average cost of using the method (including QA/QC considerations)? ▪ $20-30 per sample for chlorophyll a ▪ $80-150 per sample for generic id (cell counts only) **b)** Is the indicator readily and/or commercially available? There are a limited number of experienced algal taxonomists available and turn around time must be considered. The analysis of chlorophyll a as a measure of biomass can be more easily performed in many laboratories with the appropriate equipment.

5.25 References[1]

Abbott, R.T. 1974. American seashells. (2nd Ed.) Van Nostrand Reinhold Co., New York. 633 pp.

AETE (Aquatic Effects Technology Evaluation Program). 1996. 1995 Field evaluation of aquatic effects monitoring methods-Pilot study, Volume 1, AETE Project 4.1.1. Canada Centre for Mineral and Energy Technology, Ottawa, ON.

Aquatic Effects Technical Evaluation (AETE) Committee. 1995. Pilot Field Site: Field Trip Report and Recommendations.

Aquatic Effects Technical Evaluation (AETE) Management Committee. (1995) AETE program:1995 approved work plan. Canada centre for Mineral and Energy Technology (CANMET), Natural Resources Canada, Ottawa, Ontario, 25pp.

Alberta Environment. 1990. Selected Methods for the Monitoring of Benthic Invertebrates in Alberta Rivers. Environmental Quality Monitoring Branch, Alberta Environment.

Alldredge, J.R. 1987. Sample size for monitoring of toxic chemical sites. Envir. Monitor. Assess. 9:143-154.

Appy, T.D., I.F. Linkletter, and M.J. Dadswell. 1980. A guide to the marine flora and fauna of the Bay of Fundy: Annelida: Polychaeta. Fishereies and Marine Service, Technical Report No. 920. 124 pp.

Austin, W.C., 1985. An Annotated Checklist of Marine Invertebrates in the Cold Temperate Northeast Pacific, Vol. 1, 2 and 3, Khoyatan Marine Laboratory, Cowichan Bay, B.C., Contract to Fisheries and Oceans.

Austin A. 1983. Evaluation of changes in a large oligotrophic wilderness park lake exposed to mine tailing effluent for 14 years: The periphyton. Natur. Can.110: 119-134.

Austin A. and J. Deniseger. 1985. Periphyton community changes along a heavy metal gradient in a long narrow lake. Environ. Exp. Bot. 25: 41-52.

Austin A. and N. Munteanu. 1984. Evaluation of changes in a large oligotrophic wilderness park lake exposed to mine tailing effluent for 14 years: the phytoplankton. Environ. Pollut. 33: 39-62

Bailey, R.C., Norris R.H., and T.B. Reynoldson. (eds.) 1996. Study design and data analysis in benthic macroinvertebrate assessments of freshwater ecosystems using a

---

[1] List incomplete. Please contact the National EEM Office if you require a missing reference (http://www.ec.gc.ca/eem).

reference site approach. Technical Information Workshop, North American Benthological Society, 44th Annual Meeting, Kalispell, Montana. 69p.

Baker, H.R. 1980. Key to the Common Tubificid Species of the Northeast Pacific, Manuscript from Oligochaeta Workshop.

Banse, K. 1972. Redescription of some species of Chone Kroyer and Euchone Malmgren, and three new species (Polychaeta: Sabellidae) Fishery Bulletin 70(2):459-495.

Barber, W.E., and N.R. Kevern. 1974. Seasonal variation of sieving efficiency in a lotic habitat. Freshwat. Biol. 4:293-300.

Beak 1999. Quality Assurance program for assessing mine-related effects using benthic invertebrate communities. AETE project 2.1.4 Natural Resources Canada, CANMET, Ottawa ON, 33 pp.

Bednarik, A.F., and W.P. McCafferty. 1979. Biosystematic revision of the genus Stenonema. Can. Fish. Aquat. Sci. Bull. No. 201, 73 p.

Beyers, D.W. 1998. Causal inference in environmental impact studies. J.N. Am. Benthol. Soc. 17:367-373.

Beyers, R. J. and H.T. Odum.1993. Ecological microcosms. Springer-Verlag, New York, NY, USA.

Blake, J.A. 1971. Revision of the genus Polydora from the east coast of North America (Polychaeta: Spionidae). Smithson. Contr. Zool. 75:1-32.

Blake, J.A. 1991. Revision of some genera and species of Cirratulidae (Polychaeta) from the western North Atlantic. Ophelia Suppl. 5:17-30.

Bloom, S.A. 1981. Similarity indices in community studies: Potential pitfalls. Mar. Ecol. Prog. Ser. 5:125-128.

Booth, J., Hay, D. and Truscott, J. 1996. Standard methods for sampling resources and habitats in coastal subtidal regions of British Columbia: Part I - Review of mapping with preliminary recommendations. Can. Tech. Rep. Fish. Aquat. Sci. 2118. 53pp.

Bousfield, E.L. 1958. Freshwater amphipod crustaceans of glaciated North America. Can. Field Nat. 72(2): April-June

Bousfield, E.L. 1960. Canadian Atlantic seashells. Canada Department of Northern Affairs and Natural Resources, National Museum of Canada. 72 pp.

Bousfield, E.L. and E.A. Hendryks, 1994. The Amphiphod Superfamily Leucothoidea on the Pacific Coast of North America. Family Pleustidae: Subfamily Pleustinae, Systematics and Biogeography, Amphipacific Vol.1, No. 2.

Bousfield, E.L. and E.A. Hendryks. 1995a. The Amphipod Family Pleustidae on the Pacific Coast of North America. Part III, Subfamilies Parapleustinae, Dactylopleustinae,and Pleusirinae: Systematics and Distributional Ecology, Amphipacifica Vol.II, No.1.

Bousfield, E.L. and E.A. Hendryks.1995b. The Amphipod Superfamily Eusiroidea in the North Pacific Region.I Eusiridae: systematics and distributional ecology, Amphipacifica Volume 1, Number 4.

Bousfield, E.L. and P.M. Hoover. 1995. The Amphipod Superfamily Pontoporeioidea on the Pacific Coast of North America. II. Family Haustoriidae. Genus Eohaustorias J.L. Barnard: Systematics and Distributional Ecology, Amphipacifica, Vol. II, No. 1.

Bousfield, E.L., and J. A. Kendall. 1994. The Amphipod Superfamily Dexaminoidea of the North American Pacific Coast; Families Atylidae and Dexaminidae: Systematics and Distributional Ecology, Amphipacifica Vol. 1, No. 3.

Brandlova, J. 1972. The cladocera of Ontario with remarks on some species and distribution. Can. J. Zool. 50: 1373-1403.

Brinkhurst, R. O. 1982. British and other marine and estuarine oligochaetes. Synopses of the British Fauna, N.S. 21: 1-127.

Brinkhurst, R.O. 1986. Guide to the freshwater aquatic microdrile Oligochaetes of North America. Can. Spec. Publ. Fish. Aquat. Sci. 84. 259 p.

Burch, J.B. 1975a. Freshwater Sphaeriacean clams (Mollusca: Pelecypoda) of North America, Malacological Publications, Hamburg, Michigan.

Burch, J.B., 1975b. Freshwater Unionacean Clams (Mollusca: Pelecypoda) of North America, Malacological Publications, Hamburg, Michigan
Burd, B.J., A. Nemec, and R.O. Brinkhurst. 1990. The development and application of analytical methods in benthic marine infaunal studies. Adv. Mar. Biol. 26:169-247.

Butler, T.H..1983. Shrimps of the Pacific Coast of Canada, Bulletin 202, Dept. of Fish. And Oceans, Ottawa.

Cain J.R., D.C. Paschal and C.M. Hayden. 1980. Toxicity and bioaccumulation of cadmium in the colonial green alga *Scenedesmus obliquus*. Arch. Environ. Contam. Toxicol. 9: 9-16

Camburn, K.E., J.C. Kingston and D.F. Charles (eds.). 1984-1986. PIRLA Diatom Iconograph. Contains 53 photographic plates and 1059 figures, plus figure legends. PIRLA unpublished Report Series, Report No. 3. Indiana University, Bloomington, Indiana.

Canton, L.W. 1991. Improved subsampling methods for the EPA "Rapid Bioassessment" benthic protocols. Bull. N. Am. Ben. Soc. 8:317-319.

Cattaneo, A., and J. Kalff. 1980. The relative contribution of aquatic macro-phytes and their epiphytes to the production of macrophyte beds. Limnol. Oceanogr. 25: 280-289

Chessman B.C. 1985. Artificial–substratum periphyton and water quality in the lower La Trobe River, Victoria. Aust. J. Mar. Freshwater Res. 36: 855-871

Chu, H.F. and L.K. Cutkomp, 1992. How to Know the Immature Insects, Wm. C. Brown Publishers, Dubuque, Iowa, 2nd edition.

Clarke, A. H. 1981. The freshwater molluscs of Canada. National Museum of Natural Sciences. Ottawa, Ontario. 446p.

Clarke, K.R. 1990. Comparison of dominance curves. Mar. Biol. Ecol. 138:143-157.

Clark, K.R. 1993. Non-parametric multivariate analyses of changes in community structure. Aust.J.Ecol., 18:117-143.

Clarke, K.R. and M. Ainsworth. 1993. A method of linking multivariate community structure to environmental variables. Mar.Ecol.Prog.Ser., 92:205-219.

Clarke, K.R., and R.H. Green. 1988. Statistical design and analysis for a 'biological effects' study. Mar. Ecol. Prog. Ser. 46:213-226.

Clements W.H. 1991. Community responses of stream organisms to heavy metals: a review of observational and experimental approaches. *In* Metal ecotoxicology. Concepts and applications. *Edited by* M. C. Newman and A. W. McIntosh. Lewis Publ. Chelsea, Michigan. pp. 363-391

Clements W.H. and P.M. Kiffney. 1994. Assessing contaminant effects at higher levels of biological organization. Environ. Toxicol. Chem. 13:357-359

Clements W.H. and D.M. Kiffney 1994. Integrated laboratory and field approach for assessing impacts of heavy metals at the Arkanas River, Colorado. Environ. Toxicol. Chem. 13: 397-404

Clifford, H.F. 1991. Aquatic Invertebrates of Alberta. The University of Alberta Press. Edmonton, Alberta. 538p.

Coates, K. A.,1980. Keys to Intertidal Genera and species of Enchytraeidae found in British Columbia, Manuscript from Oligochaeta Workshop, U. Vic.

Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Lawrence Erlbaum Associates, Hillsdale, NJ, 567pp.

Conquest, L.L., S.C. Ralph, and R.J. Naiman. 1994. Implementation of large-scale stream monitoring efforts: sampling design and data analysis issues. Pages 69-90 in S.L. Loeb and A. Spacie (editors). Biological monitoring of aquatic systems. Lewis Publishers, Ann Arbor Michigan.

Cooper, S.D., Barmuta L. 1993. Field experiments in biomonitoring. In D.M. Rosenberg and V.H. Resh, eds., Biomonitoring and benthic invertebrates. Chapman and Hall, New York, NY, USA, pp. 399-441.

Corkum, L.D. and Currie, D.C. 1987. Distributional patterns of immature Simuliidae (Diptera) in northwestern North America. Freshwat.Biol., 17:201-221

Corkum L.D. 1989. Patterns of benthic invertebrate assemblages in rivers of northwestern North America . Freshwater Biology 21: 191-205..

Corkum, L.D. 1991. Spatial patterns of macroinvertebrate distributions along rivers in eastern deciduous forest and grassland biomes. J.N.Am.Benthol.Soc., 10: 358-371.

Corkum L.D. 1992. Spatial distributional patterns of macroinvertebrates along rivers within and among biomes. Hydrobiologia 239:101-114.

Cowardin, L.M., V. Carter, G.C. Golet, and E.T. LaRoe. 1979. Classification of wetlands and deepwater habitats of the United States. U.S. Fish and Wildlife Service, FWS/OBS-79/31.

Cowardin, L.M. and Golet, C.F. 1995. US Fish and Wildlife Service 1979 Wetland Classification: a Review. Vegetatio 118: 139-152.

Crisp, D.J. 1984. Energy flow measurements, pp. 284-370 in Holme, N.A., McIntyre, A.D. (eds). Methods for the study of marine benthos. IBP Handbook 16. Blackwell Scientific Publications. 387 pp.

Cuffney, T.F., M.E. Gurtz, and M.R. Meador. 1993. Guidelines for the processing and quality assurance of benthic invertebrate samples collected as part of the national water-quality assessment program. U.S. Geological Survey Open-File Report 93-407.

Culp, J.M. 1999. Weight-of-Evidence Assessments for Large Rivers: Linking Community Bioassays with Field and Laboratory Studies. Workshop on Integrated Approaches for Interpreting Environmental Effects Monitoring Data. DOE EEM/1999/1.

Culp, J.M., R.B. Lowell and K.J. Cash. 2000. Integrating in situ community experiments with field studies to generate weight-of-evidence risk assessments for large rivers. Environ. Toxicol. Chem. 19(4):(in press).

Culp J.M, Podemski C.L, Cash K.J, Lowell R.B. 1996. Utility of field-based artificial streams for assessing effluent effects on riverine ecosystems. J. Aquat. Ecosystem Health 5:117-124.

Culp J.M, Lowell R.B, Cash K.J. 2000a. Integrating mesocosm experiments with field and laboratory studies to generate weight-of-evidence risk assessments for large rivers. Environ. Toxicol. Chem. 19:1167-1173.

Culp J.M., Podemski C.L., Cash K.J., Lowell R.B. 2000b. A research strategy for using stream microcosms in ecotoxicology: integrating experiments at different levels of biological organization with field data. J. Aquat. Ecosystem Stress and Recovery 7:167-176.

Cutler, E.B. 1973. Sipuncula of the Western North Atlantic. Bull. Am. Mus. Nat. Hist. 152(3):1-204.

Deniseger J., L. J. Erickson, A. Austin, M. Roch and M. J. R. Clark. 1990. The effects of decreasing heavy metal concentrations on the biota of Buttle Lake, Vancouver Island, British Columbia. Wat. Res. 24:403-416

Downing, J.A. 1979. Aggregation, transformation and the design of benthos sampling programs. J. Fish. Res. Bd Can. 36: 1454-1463.

Downing , J.A. 1981. How well does the fourth-root transformation work? Can. J. Fish. Aquat. Sci. 38: 127-129.

Downing, J.A. 1984. Sampling the benthos of standing waters. In A Manual on Methods for the Assessment of Secondary Productivity in Fresh Waters, Second Edition (J.A. Downing and F.H. Rigler, eds),. Blackwell Scientific, Oxford, pp. 87-130. IBP Handbook 17.

Downing, J.A. 1986. Spatial heterogeneity: Evolved behaviour or mathematical artefact? Nature, London 323: 255-257.

Dubé M.G., Culp J.M., Smith J.A. 1998. Development and testing of a mesocosm technology for environmental effects monitoring (EEM) methods application. pp. 43-46. In S.C. Courtenaay, W.R. Parker and G. P. Rawn [eds.] Proceedings of a workshop to assess alternatives to the fish survey component of the environmental effects monitoring program for Canadian pulp and paper mills. Can. Tech. Rep. Fish. Aquat. Sci. No. 2233.

Dussart, B. 1969. Les copépodes des eaux continentales. N. Boubee & Cie (ed.) 292p.

Edmunds, G.F. Jr., S.L. Jensen, and L. Berner. 1976. The mayflies of North and Central America. University of Minnesota Press. Minneapolis. 330p.

Eleftheriou, A. and Holme, N.A. 1984. Macrofauna techniques; pp. 140-216 in; Holme, N.A., McIntyre, A.D. (eds). Methods for the study of marine benthos. IBP Handbook 16. Blackwell Scientific Publications. 387 pp.

Elliott, J.M. 1977. Some methods for the statistical analysis of samples of benthic invertebrates. Freshwater Biological Assoc. Sci Publ. No. 25 160 pp.

Elliott, J.M., and P.A. Tullett. 1978. A bibliography of samplers for benthic invertebrates. Freshwater Biological Association, Occasional Publication, no. 4, 61 pp.

Elliott, J.M., and P.A. Tullett. 1983. A supplement to a bibliography of samplers for benthic macroinvertebrates. Freshwater Biological Association, Occasional Publication, no. 20, 27 pp.

Environment Canada & Department of Fisheries and Oceans. 1993. Technical guidance document for aquatic environmental effects monitoring related to federal *Fisheries Act* requirements. Version 1.0. Environment Canada, April 1993, 128 pp.

Environment Canada, and Department of Fisheries and Oceans. 1995. Further guidance for the invertebrate community survey for aquatic environmental effects monitoring related to federal *Fisheries Act* requirements. Report No. EEM 2. Evaluation and Interpretation Branch, Environment Canada, Ottawa, ON, 206pp.

Environment Canada. 1998. Pulp and Paper Technical Guidance for Aquatic Environmental Effects Monitoring. EEM/1998/1. Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch, Ottawa, ON, Canada.
Environment Canada 1999. "Biological Test Method: Test for Measuring the Inhibition of Growth Using the Freshwater Macrophyte Lemna minor", Conservation and Protection, Ottawa, Ontario, Report EPS 1/RM/37, March 1999.

Environment Canada 1999. Metal Mining Environment Effects Monitoring Requirements. EEM/1999/1.

Faith, D.P., Humphrey, C.L. and Dostine, P.L. 1991. Statistical power and BACI design in biological monitoring. Comparative evaluation of measures of community dissimilarity based on benthic macroinvertebrates communities in Rockhole Mine Creek, Northern Territory, Australia. Aust J. Mar. freshwat. Res. 42:589-602.

Feltmate, B.W. and Fraser, B.G. 1998. Determining mining related impacts utilizing benthic macroinvertebrates population level fitness parameters. Canada Centre for Mineral Energy Technology (CANMET) Mining Association of Canada on behalf of Aquatic Effects Technology Evaluation (AETE) Program (AETE Project 2.1.5) July 1998.

Finlayson, C.M. and van der Valk, A.G. 1995. Wetland classification and inventory. A summary. Vegetatio 118: 1-2: 185-192.

Fitzpatrick, J. F. Jr., 1983. How to Know the Freshwater Crustacea, Wm. C. Brown Publishers, Dubuque, Iowa.

Fournier, J.A., and M.E. Petersen. 1991. *Cossura longocirrata* - Redescription and distribution, with notes on reproductive biology and a comparison of described species of Cossura (Polychaeta: Cossuridae). Ophelia Suppl. 5:63-80.

Fox, G.A. 1991. Practical causal inference for ecoepidemiologists. J. Toxicol. Environ. Health 33:359-379.

Frith, H.R., Seraring, G., Wainwright, P., Harper, H. and Emmett, B. 1993. Review of habitat classification systems and an assessment of their suitability to coastl B.C. Unpub. report to Environmental Emergency services Branch, Env. Can. from L.G.L. Ltd., Sidney, B.C.

Gibbons, W.N., M.D. Munn, and M.D. Paine. 1993. Guidelines for monitoring benthos in freshwater environments. Prepared for Environment Canada, North Vancouver, B.C. by EVS Consultants, North Vancouver, B.C. 81 pp.

Gilbertson, M. 1997. Advances in forensic toxicology for establishing causality between Great Lakes epizootics and specific persistent toxic chemicals. Environ. Toxicol. Chem. 16:1771-1778.

Glozier, N. 1989. The effects of biotic and abiotic factors on the foraging success of a lotic minnow, Rhinichthys cataractae. M.Sc. Thesis, 151 p.

Gosner, K.I. 1971. Guide to the identification of marine and estuarine invertebrates. J. Wiley and Sons. New York. 693 pp.

Graney F.L., Kennedy J.J., Rodgers J.H. (eds.) 1994. Aquatic mesocosm studies in ecological risk assessment. Special publication of SETAC. Press, Pensacola, FL, USA.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley and Sons, Toronto. 257 pp.

Green, R.H. 1989. Power analysis and practical strategies for environmental monitoring. Envir. Res. 50:195-205.

Green, G. 1994. Protocols for reference and voucher collections of aquatic invertebrates stored at the Royal British Columbia Museum. DOE-FRAP 1994-15.

Guckert J.B. 1993. Artificial streams in ecotoxicology. *In* G. A. Lamberti and A. D. Steinman, eds., Research in artificial streams: applications, uses, and abuses, J. North Amer. Benthological Soc. 12:313-384, pp. 350-356.Guenther, W. C. 1981. Sample size formulas for normal theory *T* tests. Am. Stat. 35:243-244.

Hargrave, B.T., Duplisea, D.E., Pfieffer, E. and D.J. Wildish. 1993. Seasonal changes in benthic fluxes of dissolved oxygen and ammonia associated with marine cultured Atlantic salmon. Mar.Ecol.Prog.Ser. 96: 249-257.

Hargrave, B.T., Phillips, G.A., Doucette, L.I., White, M.J., Milligan, T.G., Wildish, D.J. and R.E. Cranston. 1997. Assessing benthic impacts of organic enrichment from marine aquaculture. Wat.Air Soil Pollut. 99:641-650.

Hart, J.F.L.1982. Crabs and Their Relatives of British Columbia, Handbook No. 40, B.C. Provincial Museum.

Hauer, F.R., and Lamberti, G.A. (editors). 1996. Methods in stream ecology. Academic Press Inc., San Diego California. 674p.

Hay, D.E., Waters, R.D. and Boxwell (eds.) 1996. Proceedings Marine Ecosystem Monitoring Network Workshop, Nanaimo, B.C. March 28-30. 1995. Can. Tech. Rep. Fish. Aquat. Sci. 2108. 160pp.

Hensler, G.L. 1984. An introduction to statistical and ecological software. In: Statistics in the Environmental Sciences, Gertz, S.M. and M.D. London (eds.), STP 845, American Society for Testing and Materials, Philadelphia,

Hellawell, J.M. (1986) Biological Indicators of Freshwater Pollution and Environmental Management. Elsevier Applied Science Publishers, London. 546p.
Hickley, P., 1975. An apparatus for subdividing benthos samples. Oikos 26:92-96.

Hoaglin, D.C., Mosteller, F. and Tukey, J.W. 1983. Understanding robust and exporatory data analysis. John Wiley, New York.

Hobson, K.D. and K. Banse. 1981. Sedentariate and Archiannelid Polychaetes of British Columbia and Washington, Bulletin 209, Ottawa.

Holme, N.A., McIntyre, A.D. (eds). 1984. Methods for the study of marine benthos. IBP Handbook 16. Blackwell Scientific Publications. 387 pp.

Howmiller, R.P. 1972. Effects of preservative on weights of some common macrobenthic invertebrates. Trans. Amer. Fish. Soc. 101:743-746.

Hoyle, M.H. 1973. Transformations - an introduction and a bibliography. Inter. Stat. Rev. 41:203-223.

Hughes R.M. 1995. Defining acceptable biological status by comparing with reference conditions. pp 31-47 In Biological Assessment and Criteria: tools for water resource planning and decision making Edited by Davis, W.S. and Simon TP Lewis Publisher, Boca Raton 415p.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monog. 54:187-211.

Jackson, D.A 1993. Multivariate analysis of benthic invertebrate communities: the implication of choosing particular data standardization's, measures of association, and ordination measures. Hydrobiologia 268:9-26.

Johannsen, O.A., 1977. Aquatic Diptera, Entomological Reprint Specialists, LosAngeles, Calif., fourth reprinting.

Johnson, R.K., Wiederholm, T., and L. Eriksson. 1991. The influence of season on the classification and ordination of profundal communities of nutrient poor, oligo-mesohumic Swedish lakes using environmental data. Verh.Internat.Verein.Limnol.

Johnson, R.K., T. Wiederholm, and D.M. Rosenberg. 1993. Freshwater biomonitoring using individual organisms, populations, and species assemblages of benthic macroinvertebrates. In Freshwater Biomonitoring and Benthic Macroinvertebrates (D.M. Rosenberg, and V.H. Resh , eds),. pp 40-125.

Jonasson, P.M. 1955. The efficiency of sieving techniques for sampling freshwater bottom fauna. Oikos 6:183-207.

Keen, A.M. and E. Coan. 1974. Marine molluscan genera of western North America. Stanford University Press.

Keough, M.H. and B.D. Mapstone. 1995. Protocols for designing marine ecological monitoring programs associated with BEK mills. National Pulp Mills Research Program. Technical Report No. 11.

Kilgour, B.W. 1997. Ph.D. Thesis (in prep.). Biology Department, University of Waterloo, Waterloo, ON.

Kilgour, B.W. 1997. Fish-benthos correlation's and effects on benthos that reflect significant effects on fish communities in southern Ontario streams. PhD thesis, University of Waterloo. 190pp.

Klemm, D.J. 1972. Freshwater leeches (Annelida: Hirudinea) of North America. U.S. Environmental Protection Agency, Identification Manual no 8., 53 p.

Klemm, D.J. 1985. A guide to the Freshwater Annelida (Polychaeta, Naidid and Tubificid Oligochaeta, and Hirudinea) of North America, Kendall/Hunt Publ. Co. Dubuque, Iowa, 198 p.

Klemm, D.J. 1991. Taxonomy and pollution ecology of the Great Lakes region leeches (Annelida: Hirudinea). Michigan Academician. 24: 37-103.

Klemm, D.J., P.A. Lewis, F. Fulk, and J.M. Lazorchak. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters. EPA 600/4-90/030. U.S. Environmental Protection Agency, Environmental Monitoring Laboratory, Cincinatti, Ohio.

Kosinski, R.J. 1989. Artificial streams in ecotoxicological research. *In* A. Boudou and F. Ribeyre, eds., Aquatic ecotoxicology: fundamental concepts and methodologies, Vol. I, CRC Press, Boca Raton, FL, USA, pp. 297-316.

Kozloff, E.N. 1987. Marine Invertebrates of the Pacific Northwest, Univ. of Wash. Press.

Kreis, R.G., Jr. 1986. Variability study. Section 17. In Paleoecological Investigation of Recent Lake Acidification (PIRLA): Methods and project description (Charles, D.F. and D.R. Whitehead, eds). Electric Power Research Institute, Palo Alto, California.
Kreis, R.G., Jr. 1989. Variability Study - Interim Results. Section 4. In Paleoecological Investigation of Recent Lake Acidification (PIRLA): 1983-1985: Interim Report (Charles, D.F. and D.R. Whitehead, eds.). October 1989. Electric Power Research Institute, Palo Alto, California

Kronberg, I. 1987. Accuracy of species and abundance minimal areas determined by similarity area curves. Mar. Biol. 96: 555-561.

Lambert, P. 1981. The sea stars of British Columbia, B.C. Provincial Museum Handbook No. 39.

Laplante, S., Y. Bousquet, P. Bélanger, & C. Chantal. 1991. Liste des espèces de Coléoptères du Québec. Fabreries, Supplément 6. 136 p.

Laubitz, D.R. 1972. The Caprellidae (Crustacea, Amphipoda) of Atlantic and Arctic Canada. National Museum of Canada, Publications in Biological Oceanography, No. 4. 82 pp.

Lehmkuhl, D.M. 1975a. Field guide to aquatic insect families. Blue Jay 33:199-219.

Lehmkuhl, D.M. 1975b. Saskatchewan damselflies and dragonflies. Blue Jay 33:18-27.

Lehmkuhl, D.M. 1976. Mayflies. Blue Jay 34:70-81.

Lehmkuhl, D.M. 1979. How to know the aquatic insects. Wm. C. Brown Company Publishers. Dubuque, Iowa.168p.

Leopold, L.B., Wolman, G.M., and J.P. Miller. 1964. Fluvial processes in geomorphology. W.H. Freeman and Co., San Francisco. 522 p.

Leopold, L.B., 1994. A veiw of the river. Harvard University Press, Cambridge. 297 p.

Lenat D, Resh VH. 2000. Taxonomy, stream macroinvertebrates, and water quality assessments – what's the value of having a name? Journal of the North American Benthological Society Bridges.

Lewis, P.A. 1974. Taxonomy and ecology of Stenonema mayflies. EPA-670/4-74-006. 81 p.

Leuven, R.S.E.W., Brock, T.C.M., and H.A.M. van Druten. 1985. Effects of preservative on dry- and ash-free dry weight biomass of some common aquatic macro-invertebrates. Hydrobiologia 127:151-159.

Levings, C.D. and Thom, R.M. 1994. Habitat changes in the Georgia Basin: Implications for resource management and restoration. p 330-351 In Review of the marine environment and biota of Strait of Georgia, Puget Sound and Juan de Fuca Strait: Proceedings of the BC/Washington Symposium on the Marine Environment Jan 13,14 1994 (Wilson, R.C.H., Beamish, R.J., Aitkens, F. and Bell, J., eds). Can. Tech. Rep. Fish. Aquat. Sci. 1948.

Lewis, P.A. 1974. Taxonomy and ecology of Stenonema mayflies. EPA-670/4-74-006. 81 p.

Lowe, R.L. and Pan, Y. 1996. Benthic algal communities as biological monitors. Chapter 22 p 705-739 in "Algal Ecology". Academic Press, Inc.

Lowell, R.B, J.M. Culp and F.J. Wrona. 1995. Stimulation of increased short-term growth and development of mayflies by pulp mill effluent. *Environ. Toxicol. Chem.* 14:1529-1541.

Lowell , R.B. 1997. Discussion paper on critical effect size guidelines for EEM using benthic invertebrate communities. Report to the Environmental Effects Monitoring Program. EEM/1997/9

Lowell, R.B., J.M. Culp and M.G. Dube. 2000. A weight-of evidence approach for northern river risk assessment: integrating the effects of multiple stressors. Environ. Toxicol. Chem. 19(4):(in press).

Ludwig, J.A. and J.F. Reynolds. 1988. Statistical Ecology: A Primer on Methods and Computing. John Wiley and Sons, New York.

Malley, D.F., and J.B. Reynolds. 1979. Sampling strategies and life history of non-insectan freshwater invertebrates. J. Fish. Res. Bd. Can. 36:311-318.

Mapstone, B.D. 1995. Scalable decision criteria for environmental impact assessment: Effect size, Type I and Type II errors. In Detecting Ecological Impacts: Concepts and

Applications in Coastal Habitats (Schmitt, R.J. and C.W. Osenberg, eds.), Academic Press, San Diego. pp. 67-80.

Marchant, R.A. 1989. A subsampler for samples of benthic invertebrates. Bull. Aust. Soc. Limnol. 12:49-52.

Mason, C.F. 1991a. Biology of Freshwater Pollution, Second Edition. Longman Scientific & Technical, New York. 351 pp.

Mason, W.T. Jr. 1991b. Sieve sample splitter for benthic invertebrates. J. Freshwat. Ecol. 6:445-449.

Matthews, G.W.T. 1993. The Ramsar Convention; its history and development. Ramsar Caonvention Bureau, Glan Switzerland

Maxwell JR, Edwards CJ, Jensen ME, Paustian SJ, Parrott H, Hill DM. 1995. A hierarchical framework of aquatic ecological units in North America (Neartic Zone). Gen Tech., Rep. NC-176. St Paul, MIN. U.S. Department of Agriculture, Forest Service, North Central Forest Experimental Station. 72p.

McCafferty, W.P. and Waltz, R.D. 1990. Revisionary synopsis of the Baetidae (Ephemeroptera) of North and Middle America. Trans. Amer. Entomol. Soc. 116:769-800.

Meador, M.R., Hupp, C.R., Cuffney, T.F. and M.E. Gurtz 1993. Methods for characterizing stream habitat as part of the national water-quality assessment program. U.S. Geological Survey. Open-File Report 93-408, Raleigh, North Carolina.

Meador, M.R., Cuffney, T.F., and M.E. Gurtz 1993, Methods for sampling fish communities as a part of the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 93-104, 40p.

Merritt, R.W., and K.W. Cummins. 1984. An Introduction to the Aquatic Insects of North America, Second Edition. Kendall/Hunt, Dubuque. 722 pp.

Morihara, D.K. and W.P. McCafferty. 1979. The Baetis larvae and North America (Ephemeroptera: Baetidae). Trans. Amer. Ent. Soc. 105:139-221.

Morris, P. A. 1951. A field guide to the shells of our Atlantic and Gulf coasts. The Peterson field guide series. Houghton Mifflin Company, Boston, Mass. 236 p.

Newbury, R.W. 1984. Hydrolic determinants of aquatic insect habitats. In The Ecology of Aquatic Invertebrates (V.H. Resh and D.M. Rosenberg, eds),. Praeger Publishers, New York.

Newbury, R.W., & M. Gaboury. 1993. Stream analysis: fish habitat and design. A field manual. Newbury Hydraulics, Gibsons, B.C.

Norris, R.H., and A. Georges. 1993. Analysis and interpretation of benthic macroinvertebrate surveys. In Freshwater Biomonitoring and Benthic Invertebrates (D.M. Rosenberg and V.H. Resh, eds.) pp. 234-286.

Oliver, D.R. et al, 1978. A Key to some Larvae of the Chironomidae (Diptera) from the Mackenzie and Porcupine River Watersheds, Fisheries and Marine ServiceTechnical Report 791, Agriculture Canada, Ottawa.

Omernik, J.M. 1995. Ecoregions: a spatial framework for environmental management. Pp. 49-62. *In* Biological Assessment and Criteria. Tools for Water Resource Planning and Decision Making. *Edited by* Davis, W.S. & T.P. Simon. Lewis Publ., Boca Raton 415p.

Ormerod, S.J. and R.W. Edwards. 1987. The ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye in relation to environmental factors. Freshwat.Biol., 17:533-546.

Pearson, T.H. 1975. The benthic ecology of Loch Linnhe and Loch Eil, a sea-loch system on the west coast of Scotland. IV: Changes in the benthic fauna attributable to organic enrichment. J. Exp. Mar. Biol. Ecol. 20: 1-41.

Peckarsky, B.L., Fraissinet, P.R., Penton, M.A., and Conklin, D.J. (eds), 1990. Freshwater macroinvertebrates of northeastern North America. Cornell University Press, 442p.

Pennak, R.W. 1978. Freshwater invertebrates of the United States: Protozoa to Mollusca. 3rd ed. John Wiley & Sons Inc., New York. 803 p.

Plafkin, J.L., M.T. Barbour, M.T. and K.D. Porter. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. EPA/444/4-89-001, 162 pp.

Pontasch, B.C., Smith, E.P., and Cairns, J. Jr. 1989. Diversity indices, community comparison indices and canonical discriminant analysis: Interpreting the results of multispecies toxicity tests. Water Res. 23:1229-1238.

Pontasch K.W. 1995. The use of stream microcosms in multispecies testing. pp. 169-191 In J. Cairns and B. R. Niederlehner, eds, Ecological toxicity testing: scale, complexity and relevance, Lewis Publishers, Boca Raton, FL, USA,.

Rabeni, C.F., and K.E. Gibbs. 1978. Comparison of two methods used by divers for sampling benthic invertebrates in deep rivers. J. Fish. Res. Bd. Can. 35:332-336.

Rees, H.L. 1984. A note on mesh selection and sampling efficiency in benthic studies.

Mar. Poll. Bull. 15:225-229.

Reish, D.J. 1959. A discussion of the importance of the screen size in washing quantitative marine bottom samples. Ecology 40:307-309.

Resh, V.H., and D.M. Rosenberg. 1984. The Ecology of Aquatic Insects. Praeger Publishers.

Resh, W.H., Norris, R.H. and Barbour, M.T. 1995. Design and implementation of rapid assessment approaches for water resource monitoring using benthic macroinvertebrates. Aust. J. Ecol. 20:108-221.

Resh V.H, McElvray E.P. 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates, p. 159-194. In Rosenberg D.M, Resh V.H (ed.), Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

Reynoldson, T. B., R. C. Bailey, K. E. Day, and R. H. Norris. 1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology 20:198-219.

Reynoldson, T.B. and D.M. Rosenberg. 1996. Sampling strategies and practical considerations in building reference data bases for the prediction of invertebrate community structure. Pp 1-31. In Study design and data analysis in benthic macroinvertebrate assessments of freshwater ecosystems using a reference site approach (Bailey, R.C., Norris R.H., and T.B. Reynoldson. , eds.) Technical Information Workshop, North American Benthological Society, 44th Annual Meeting, Kalispell, Montana. 69p.

Reynoldson, T.B., R.H. Norris, V.H. Resh, and D.M. Rosenberg. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society 16: 833-852

Ripley E.A., R.E. Redman and A.A. Crowder. 1996. Environmental effects of mining. St. Lucie Press, Delray Beach, Florida. 356p.

Robinson, C.L.K. and Levings, C.D. 1995. An overview of habitat classification systems, ecological models and geographic information systems applied to shallow foreshore marine habitats. Can. Man. Rep. Fish. Aquat. Sci. 2322: 65pp.

Robinson, C.L.K., Hay, D.E., Booth, J. and Truscott, J. 1996. Standard methods for sampling resources and habitats in coastal subtidal regions of British Columbia: Part 2 - Review of sampling with preliminary recommendations. Can. Tech. Rep. Fish. Aquat. Sci. 2119. 119pp.

Roch, M., R.N. Nordin, A. Austin, C.P.J. Mckean, J. Deniseger, R.D. Kathman, J.A. McCarter, and M.J.R. Clark. 1985. The effects of heavy metal contamination on the aquatic biota of Buttle Lake and the Campbell River drainage (Canada). Arch. Environ. Contam. Toxicol. 14:347-362

Rosenberg, D.M. 1978. Practical sampling of freshwater macrozoobenthos: a bibliography of useful texts, reviews, and recent papers. Canadian Fisheries and Marine Service Technical Report 790, 15 pp.

Rosenberg, D.M., and V.H. Resh (eds). 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman & Hall, New York. 488 pp.

Saether, Ole A., 1975. Nearctic and Palaearctic Heterotrissocladius (Diptera: Chironomidae_ Bulletin of the Fisheries Research Board of Canada, Bulletin 193, Ottawa.

Saether, Ole A., 1977. Taxonomic studies on Chironomidae: Nanocladius, Pseudochironomus and the Harnischia complex. Bulletin of the Fisheries Research Board of Canada, Bulletin 196, Ottawa.

Sand-Jensen K. and L. Rasmussen. 1978. Macrophytes and chemistry of acidic streams from lignite mining areas. Bot. Tidsskrift. 72: 105-112

Santa Barbara Museum of Natural History (SBMNH). 1994a. Taxonomic Atlas, Vol. 1, Platyhelminthes, and Nemertea, Santa Barbara, California, 1994

Santa Barbara Museum of Natural History (SBMNH). 1994b. Taxonomic Atlas, Vol. 2, Porifera.

Santa Barbara Museum of Natural History (SBMNH). 1994c. Taxonomic Atlas, Vol. 4, Oligochaeta and Polychaeta: Phyllodocidae to Paralacydoniidae.

Santa Barbara Museum of Natural History (SBMNH) 1995a. Taxonomic Atlas, Vol.5, The Annelida, Part 2 Polychaeta: Phyllodocida (Syllidae and Scale-Bearing Families), Amphinomida and Eunicida.

Santa Barbara Museum of Natural History (SBMNH). 1995b. Taxonomic Atlas, Vol. 12, The Crustacea, Part 3, The Amphipoda.
Santa Barbara Museum of Natural History (SBMNH). 1995c. Taxonomic Atlas, Vol. 13, The Bryozoa,

Santa Barbara Museum of Natural History (SBMNH). 1996a. Taxonomic Atlas, Vol. 6, The Annelida, Part 3 Polychaeta: Orbiniidae to Cossuridae.

Santa Barbara Museum of Natural History (SBMNH). 1996b. Taxonomic Atlas, Vol. 9, The Mollusca, Part 2 The Gastropoda.

Santa Barbara Museum of Natural History (SBMNH). 1996c. Taxonomic Atlas, Vol. 14, Miscellaneous Taxa.

Santa Barbara Museum of Natural History (SBMNH). 1997a. Taxonomic Atlas, Vol 10, The Pycnogonida, The Crustacea Part 1 The Decapoda and Mysidacea, 1997

Santa Barbara Museum of Natural History (SBMNH). 1997b. Taxonomic Atlas, Vol. 11, The Crustacea Part 2, The Isopoda, Cumacea and Tanaidacea.

Schefter P. and G.B. Wiggins, 1986. A Systematic Study of the Nearctic Larvae of the Hydropsyche morosa Group (Trichoptera: Hydropsychidae) Royal Ontario Museum, Toronto.

Schindler, D.W. 1987. Detecting ecosystem responses to anthropogenic stress. Can J. Fish. Aquat. Sci. 44 (Suppl. 1): 6-25.

Schultz, G.A., 1969. The Marine Isopod Crustaceans, Wm. C. Brown Co., Dubuque, Ia.

Schuster, G.A. and D.A. Etnier, 1978. A Manual for the Identification of the larvae of the Caddisfly Genera Hydropsyche Pictet and Symphitopsyche Ulmer in Eastern and Central North America, U.S. Environmental Protection Agency, Cincinnati, Ohio.

Schwinghamer, P. 1981. Characteristic size distribution s of integral benthic communities. Can. J. Fish. Aquat. Sci. 38: 1255-1263.

Schwinghamer, P. 1983. Generating ecological hypotheses from biomass spectra using causal analysis: A benthic example. Mar. Ecol. Progr. Ser. 13: 151-166.

Scrimgeour, G.J., J.M. Culp and N.E. Glozier. 1993. An improved technique for sampling lotic invertebrates. Hydrobiologia 254:65-71

Scott, D.A. and Jones, T.A. 1995. Classification and inventory of wetlands. Vegetatio 118: 1-16.

Sebastien, R.J., Rosenberg, D.M. and A.P. Wiens. 1988. A method for subsampling unsorted benthic macroinvertebrates by weight. Hydrobiologia 157:69-75.

Shriner, C. and T. Gregory. 1984. Use of artificial streams for toxicological research. Crit. Rev. Toxicol. 13:353-281.

Simon T.P. 1991. Development of biotic integrity expectations for the ecoregions of Indiana. I. Central Corn Belt Planing. U.S. EPA, Region V, Environmental Sciences Division, Monitoring and Quality Assurance Branch. Ambient Monitoring Section. Chicago Il. EPA 905/9-91/025

Slack, K.V., R.C. Averett, P.E. Geeson, and R.G. Lipscomb. 1973. Methods for collection and analysis of aquatic biological and microbiological samples. In Techniques of Water-Resources Investigations of the United States Geological Survey, Chapter 4A, Book 5, pp. 1-165. U.S. Department of the Interior, Geological Survey, Washington, D.C.

Small A.M., W.H. Adey, S.M. Lutz, E.G. Reese and D.L. Roberts. 1996. A macrophyte-based rapid biosurvey of stream water quality: restoration at the watershed scale. Restoration Ecol. 4:124-145

Snelgrove, R.V.R. and C.A. Butman. 1994. Animal sediment relationships revisited: cause versus effect. Oceanography and Marine Biology annual review 32:111-178.

Sortkjaer O. 1984. Macrophytes and macrophyte communities as test systems in ecotoxicological studies of aquatic systems. Ecol. Bull. 36: 75-80

Squires, H.J. 1990. Decapod crustacea of the Atlantic coast of Canada. Department of Fisheries and Oceans. Ottawa. 532 p.

St-Cyr L. Cattaneo A, Chasse R, Fraikin,C.G.J. 1997. Technical evaluation of monitoring methods using macrophytes, phytoplankton and periphyton to assess the impacts of mine effluents on the aquatic environment. AETE project 2.3.2, Natural Resources Canada, CANMET, Ottawa ON, 200 pp.

Steele, D.H., and P. Brunel. 1968. Amphipoda of the Atlantic and Arctic coasts of North America: Anonyx (Lysianassidae). J. Fish. Res. Bd. Can.. 25(5).

Stephenson, W. and Cook, S.D. 1980. Skewness of data in the analysis of species-in-sites-in-times. Proc. Royal Soc. Queensland 91:37-52.

Stewart, K.W. and B.P. Stark. 1993. Nymphs of North American Stonefly Genera (Plecoptera). University of North Texas Press, Denton, Texas. 460p

Stokes P. M. 1983. Responses of freshwater algae to metals. *In* Progress in Phycological Research, Vol.2. *Edited by* F. E. Round and A. Champman. Elsevier Science Publ. pp. 87-112.

Suess, M.J., ed. 1982 Examination of Water For Pollution Control. A Reference Handbook. Vol. 3. Biological, Bacteriological and Virological Examination. Permagon Press, Oxford, England.

Suter, G.W. 1993. A critique of ecosystem health concepts and indexes. Environ. Toxicol. Chem. 12: 1533-1539.

Suter, G.W. 1993. Retrospective risk assessment. Pages 311-364 in G.W. Suter (editor).

Ecological risk assessment. Lewis Publishers, Ann Arbor, Michigan.

Suter, G.W. II. 1993. Ecological Risk Assessment. Lewis Publishers, Boca Raton, FL, USA.

Takamura N., F. Kasai and M. M. Watanabe. 1989. Effects of Cu, Cd, and Zn on photosynthesis of freshwater benthic algae. J. Appl. Phycol. 1:39-52

Taylor B.R. 1997. Optimization of field and laboratory methods for benthic invertebrate monitoring, final report. AETE project 2.1.2 Natural Resources Canada, CANMET, Ottawa ON, 100 pp.

Taylor B.R, Bailey R.C. 1997. Technical evaluation on methods for benthic invertebrate data analysis and interpretation, final report. AETE project 2.1.3. Natural Resources Canada, CANMET, Ottawa ON, 26 pp.

Taylor, L.R. 1961. Aggregation, variance and the mean. Nature (London) 189:732-735.

Tetra Tech, Inc. 1986a. General QA/QC Considerations for Collecting Environmental Samples in Puget Sound. A Report prepared for U.S. EPA, Region 10, Seattle. Final Report TC-3991-04.

Tetra Tech, Inc. 1986b. Recommended Protocols for Station Positioning in Samples in Puget Sound. A Report prepared for U.S. EPA, Region 10, Seattle. Final Report TC-3090-05.

Tetra Tech, Inc. 1987. Recommended Protocols for Sampling and Analyzing Subtidal Benthic Macroinvertebrate Assemblages in Puget Sound. A Report prepared for U.S. EPA, Region 10, Seattle. Final Report TC-3991-04.

Thiel, T.H. 1975. The size structure of the deep-sea benthos. Int. Rev. ges. Hydrobiol. 60:575-606.

Thorp, J.H. and A.P. Covich, 1991. Ecology and Classification of North American Freshwater Invertebrates, Academic Press, San Diego, New York, Toronto.

Thorson, G. 1957. Bottom communities(sublitorral or shallow shelf). In Treatise on Marine Ecology and Paleoecology.J.W. Hedgpeth, ed.Vol 1, pp. 461-534.Memoirs of the Geological Society of America 67.

Tukey, J. 1977. Exploratory data analysis. Addison-Wesley, Reading, Mass. 688pp.

Underwood, A. J. 1993. The mechanics of spatially replicated sampling programs to detect environmental impacts in a variable world. Australian Journal of Ecology 18:99-116.

U.S. EPA [U.S. Environmental Protection Agency]. 1995. Volunteer stream monitoring: a methods manual. EPA 841 D95-001 April. Office of Wetland, Oceans and Watersheds, 4503F. Washington D.C. 20460.

U.S. EPA [U.S. Environmental Protection Agency]. 1995. Environmental monitoring and assessment program for surface waters: field operations and methods for measuring the ecological condition of wadeable streams. EPA 620/R-95-004. Office of Research and Development, U.S. Environmental Protection Agency, Washington D.C.

U.S. EPA [U.S. Environmental Protection Agency]. 1990. Biological criteria: national program guidance for sueface waters. EPA-440/5-90-004. U.S. Environmental Protection Agency, Washington, D.C., 57pp

Vezina, A.F. 1988. Sampling variance and the design of quantitative surveys of the marine benthos. Mar. Biol. 97: 151-155.

Walker, E.M. 1933. The nymphs of the Canadian species of *Ophiogomphus* Odonata, Gomphidae. Can. Ent. 65: 217-229.
Walker, E.M. 1953. The Odonata of Canada and Alaska. Part I, General, Part II. The Zygoptera-damselflies. Vol. I. Univ. Toronto Press, Toronto. 292 p.

Walker, E.M. 1958. The Odonata of Canada and Alaska. Anisoptera. Vol. 2. Univ. Toronto Press, Toronto. 318 p.

Wallace, N.A. 1919. The Isopoda of the Bay of Fundy. Univ. of Toronto Studies, Biol. Series No. 18. 42 pp.

Warwick, R.M. 1986. A new method for detecting pollution effects on marine benthic communities. Mar. Biol. 92: 557-562.

Warwick, R.M., and K.R. Clarke. 1993. Increased variability as a symptom of stress in marine communities. J. Exp. Mar. Biol. Ecol. 172:215-226.

Warwick, R.M. 1993. Environmental impact studies on marine communities: Pragmatical considerations. Aust.J.Ecol., 18:63-80.

Warwick. R.M., Pearson, T.H., and Ruswahyuni. 1987. Detection of pollution effects on marine macrobenthos: Further evaluation of the species abundance/biomass method. Mar. Biol. 95: 193-200.

Weber, C.I., ed. 1973. Biological Field and Laboratory Methods for Measuring the Quality of Surface Waters and Effluents. EPA-670/4-73-001. U.S. Environmental Protection Agency, Cincinnati, OH.

Weinberg, S. 1978. The minimal area problem in invertebrate communities of Mediterranean rocky substrata. Mar. Biol. 49: 33-40.

Wiederholm, T. 1980. Use of benthos in lake monitoring. J. Wat. Poll. Cont. Fed. 52:537-547.

Wiederhom, T. (ed.) 1983. Chironomidae of the holarctic region. Keys and diagnosis. Part 1 - Larvae. Ent. Scand. Suppl. No. 19, 457 p.

Wiederholm, T. (ed.) 1986. Chironomidae of the holarctic region. Keys and diagnosis. Part 2 - Pupae. Ent. Scand. Suppl. No. 28, 482 p.

Wiggins, G.B. 1996. Larve of the North American Caddisfly Genera (Trichoptera). 2nd Edition. University of Toronto Press. Toronto. Ontario. 457p.

Wood, D.M. 1963. The Black Flies (Diptera: Simuliidae) of Ontario. Part II Larval Identification with descriptions and illustrations. Proc. Entomol. Soc. Ontario.

Wright, J.F. 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. Australian Journal of Ecology 20: 181-197

Wright, J.F., Moss, D., Armitage, P.D., and M.T. Furse. 1984. A preliminary classification of running-water sites in Great Britain based on macroinvertebrate species and the prediction of community type using environmental data. Freshwat.Biol., 14: 221-256

Wrona, F.J., Culp, J.M., and R.W. Davies. 1982. Macroinvertebrate subsampling: a simplified apparatus and approach. Can.J.Fish.Aquat.Sci., 39: 1051-1054.

Yan N.D. 1979. Phytoplankton community of an acidified heavy metal contaminated lake near Sudbury, Ontario: 1973-1977. Water, Air, and Soil Poll. 11:43-45

Yan N.D., G.E. Miller, I. Wile and G.G. Hitchin. 1985. Richness of aquatic macrophyte floras of soft water lakes of differing pH and trace metal contents in Ontario, Canada. Aquat. Bot. 23:27-40

Yoder, C.O. 1989. The development and use of biological criteria for Ohio surface waters. Pages 139-146 in Proceedings: Water Quality Standards for the 21stCentury, 1-3 March, 1989, Dallas, TX. Criteria/Standards Division, U.S. Environmental Protection Agency, Washington, Dc.

Yoder, C.O. 1991. Answering some concerns about biological criteria based on experiences in Ohio. Pp. 95-104. *In* Water Quality Standards for the 21st Century. U.S. Environmental Protection Agency. Office of Water, Washington, D.C. 251p.

Table of Contents

6. *Effluent Characterization and Water Quality Monitoring* *6-1*

6.1 Introduction *6-1*
  6.1.1 Overview 6-1
  6.1.2 Sampling Frequency 6-1
  6.1.3 Parameters measured 6-2
  6.1.4 Sampling Locations 6-5
  6.1.5 Reporting 6-6

6.2 Effluent Characterization *6-7*
  6.2.1 Sampling Methods and Laboratory Analysis 6-7
  6.2.2 Methods for Determination of Thiosalts 6-7

6.3 Water Quality Monitoring *6-8*
  6.3.1 Preparation for the field 6-8
  6.3.2 Field Measurement of Water Quality Parameters 6-8
  6.3.3 Collection of Water Samples for Laboratory Analyses 6-10
  6.3.4 Sampling Handling, Storage and Analysis for Water Quality Monitoring 6-11
    6.3.4.1 Sample Handling and Preservation 6-11
    6.3.4.2 Sample Shipping and Storage 6-15
    6.3.4.3 Unique Sampling Requirements for Some Chemical Groups 6-15
    6.3.4.4 Laboratory Analyses of Samples 6-16
  6.3.5 Comparison of Water Quality Data in Exposure and Reference Areas 6-21
    6.3.5.1 Comparison of Results 6-21
    6.3.5.2 Estimation of Extent of Elevated Concentrations 6-22

6.4 Quality Assurance and Quality Control for Water Quality Monitoring *6-24*
  6.4.1 Standard Operating Procedures 6-24
  6.4.2 General Aspects of Quality Control in the Field 6-25
  6.4.3 Field Aspects of Quality Assurance 6-26
  6.4.4 Quality Assurance During Sample Handling, Shipping and Storage 6-26
  6.4.5 Use of Blanks and Duplicate Samples 6-29
  6.4.6 General Aspects of Quality Control in the Laboratory 6-31
  6.4.7 Details on Quality Control Aspects of Laboratory Analyses 6-31
    6.4.7.1 Good Laboratory Practices 6-31
    6.4.7.2 Calibration Control 6-32
    6.4.7.3 Precision 6-33
    6.4.7.4 Accuracy and Bias 6-34
    6.4.7.5 Recovery 6-34
    6.4.7.6 Detection Limits 6-35
    6.4.7.7 Data Reporting Conventions 6-36
    6.4.7.8 Analytical Precision and Accuracy 6-37
    6.4.7.9 Instrument Maintenance 6-37
  6.4.8 Quality Assurance in the Laboratory 6-38

6.4.9 Recording and Reporting of QA/QC Information 6-39
6.4.9.1 Documentation 6-39
6.4.9.2 Reporting of QA/QC Information 6-39

6.5 Effluent Characterization and Water Quality Monitoring during Investigation of Cause 6-40
6.5.1 Measurement of Dissolved Metals during Investigation of Cause 6-41
6.5.2 Measurement of Reagents and Reagent by-Products used in Processing.. 6-42
6.5.3 Metal speciation for metals of concern 6-42
6.5.3.1 Chemical Analysis of Metal Speciation 6-42
6.5.3.2 Chemical Speciation Models 6-43
6.5.4 Measurement of Chlorophyll A and/or Phaeo-pigment 6-43
6.5.5 Monitoring of Flow and Loadings in the Exposure Area 6-44

6.6 References 6-46

6.7 Deleterious Substances and pH as per Schedule 3 of the MMER 6-49

6.8 Parameters Required in Schedule 5 of MMER for Effluent Characterization and Water Quality Monitoring 6-51

6.9 Required Parameters for Water Quality Monitoring Only 6-54

6.10 Additional Site Specific Recommended Parameters for Effluent Characterization and Water Quality Monitoring 6-54
6.10.1 Parameters Recommended for Effluent Characterization Only 6-56
6.10.2 Site specific Parameters Recommended for Water Quality Monitoring Only6-57

List of Tables

Table 6-1 Analytical Parameters to be Measured for Effluent Characterization and Water Quality Monitoring 6-3
Table 6-2: Water sample handling recommendations regarding containers, preservatives and holding times for specific water quality parameters 6-14
Table 6-3: Typical checklist of bottle requirements for parameters for water quality monitoring 6-15
Table 6-4: Detection limits for methods used to determine metal concentrations in water samples 6-18
Table 6-5: Methodologies and detection limits used to determine concentrations of water quality parameters 6-19
Table 6-6: Summary of CCME water quality guidelines for freshwater aquatic life for key parameters monitored as part of the EEM program. 6-23
Table 6-7: Summary of recommended use of blanks and duplicate samples in the field and laboratory. Numbers are based on a batch of 20 samples. 6-30

List of Figure

Figure 6-1: Example of a chain-of-custody record (modified from U.S. EPA 1988) .. 6-28

6. Effluent Characterization and Water Quality Monitoring

6.1 Introduction

6.1.1 Overview

Effluent characterization and water quality monitoring are part of the Effluent and Water Quality Monitoring Studies required for EEM under Section 7 and Schedule 5 of MMER. The purpose of effluent characterization and water quality monitoring is to answer the question: "What is the estimated mine-related change in contaminant concentrations in the exposed area?". Data generated from effluent characterization and water quality monitoring are used to:

- monitor changes in mining operations and environmental conditions in the receiving environment;
- provide an indication of variability in effluent quality and temporal or seasonal trends
- provide supporting environmental variables to help interpret results from the biological monitoring (fish and benthic invertebrate community survey) and the sublethal toxicity testing

Effluent characterization is conducted by analysing a sample of effluent to provide information on the concentrations of contaminants in the mine effluent. In addition to the regulated deleterious substances and pH described in Schedule 3 of the proposed MMER, samples for effluent characterization will be analysed for other potential contaminants and supporting parameters (Table 6.1) four times per calendar year. Samples for effluent characterization will be aliquots of effluent samples collected for monitoring deleterious substances and pH under Section 12 and 13 of the MMER.

Water quality monitoring is conducted by collecting and analysing samples of water from the sampling areas to compare the concentrations of contaminants in the exposure area and the reference area. Samples for water quality monitoring will be analysed for the same parameters measured for effluent characterization (Table 6.1) as well as dissolved oxygen and temperature. In addition, samples of water will be collected from sampling areas in receiving environments where biological monitoring is completed, at the same time that biological monitoring is conducted.

6.1.2 Sampling Frequency

Effluent characterization and water quality monitoring will be conducted four times each calendar year, at least one month apart. It is recommended that, where possible, samples for effluent characterization and water quality monitoring are collected once in each calendar quarter. It is also recommended that samples for effluent and water be collected on the same day.

The first sample required to be collected for effluent characterization and water quality monitoring will be no later than six months after the mine becomes subject to section 7 of the MMER *(Schedule 5, s. 4(2) and s. 7(2))*.

The following factors should be taken into consideration to decide whenthe aliquots of effluent sample will be analyzed for effluent characterization :

- seasonal variability of effluent, based on composition and flow;
- the time of year when previous effluent samples have been collected;
- the time of year when sampling for water quality monitoring is being conducted;
- the time of year when concentrations of the contaminants are expected to be highest in the exposure area.

For water quality monitoring, the following factors should be taken into consideration to decide when water samples will be collected in the receiving environment:

- seasonal variability in water quality and flow in the exposure area;
- the time of year when concentrations in the exposure area of contaminants are expected to be highest;
- the time of year when previous water quality monitoring samples have been collected; and,
- the time of year when samples for effluent characterization are collected.

6.1.3 Parameters measured

Effluent characterization and water quality monitoring will be conducted for the parameters listed in Table 6.1, subject to the exemption described for mercury. Table 6.1 also includes optional parameters recommended on site specific basis that the owner or operator of a mine may record as additional supporting information and to conduct a more complete chemical characterization (Appendix 6.1).

Table 6-1 Analytical Parameters to be Measured for Effluent Characterization and Water Quality Monitoring

Deleterious substances and pH [1,2]	**Required Effluent Characterization and Water Quality Monitoring Parameters**[2,3]**:**	**Required Additional Water Quality Monitoring Parameters**	**Recommended Site-Specific Parameters**[6] **(optional)**
Arsenic	Aluminum	Dissolved oxygen[5]	Fluoride
Copper	Cadmium	Temperature[5]	Manganese
Lead	Iron		Selenium
Nickel	Mercury[4]		Uranium
Zinc	Molybdenum		Total phosphorus
Radium 226	Ammonia		Conductivity
Total cyanide	Nitrate		Calcium
Total suspended solids	Alkalinity		Chloride
pH	Total hardness		Magnesium
			Potassium
			Sodium
			Sulphate
			Total thiosalts[7]
			Water depth[5]
			Optical depth or transparency[5]
			Dissolved organic carbon
			Total organic carbon
			Salinity (marine and estuarine) [5]

---

1) List of parameters regulated (deleterious substances and pH) as per Schedule 3 of the MMER; concentration limits specified in the regulation (Schedule 4)
2) All concentrations are **total** values; dissolved concentrations may also be reported; effluent loading (Section 20 of MMER) will also be calculated and reported
3) List of parameters required for effluent characterization and water quality monitoring as per Schedule 5 of the MMER
4) Analysis of mercury may be discontinued if the concentration of total mercury in effluent is less than 0.10 µg/L in 12 consecutive samples of effluent
5) *In situ* measured parameters only for water quality monitoring (in receiving waters)
6) These other parameters are potential contaminants or supporting parameters; analysis is optional and may be added based on site specific historical monitoring data or geochemistry data
7) Measured only for effluent characterization: See Section 6.2.2 for recommended analytical method for thiosalts measurement

Samples for effluent characterization will be aliquots of samples collected for effluent monitoring under sections 12 and 13 of the MMER. Effluent characterization is conducted by analyzing and recording the hardness and alkalinity, and the concentrations, in total values of the following *(Schedule 5, s. 4(1))*:

- Aluminum;
- Cadmium;
- Iron;
- Mercury;
- Molybdenum;
- Ammonia; and
- Nitrate.

Note that, if the concentration of mercury is less than 0.10μg/L in 12 consecutive effluent samples then analyses for mercury may be discontinued *(Schedule 5, s. 4(3))*.

If the mine effluent is determined to be acutely lethal when tested under paragraph 14(1)(a) of the MMER, the mine is required to conduct effluent characterization as per Schedule 5, s. 4(1), on the aliquot of effluent tested for acute lethality.

The measurements of pH, and the concentrations of the deleterious substances (arsenic, copper, total cyanide, lead, nickel, zinc, radium 226, and total suspended solids) will be reported as part of the effluent monitoring from each mine submitted under section 21 of the MMER.

In addition to the deleterious substances, pH and the required parameters for effluent characterization, additional water quality monitoring will include the measurement of temperature and dissolved oxygen *(Schedule 5, s . 7(1)(b-d))*. Other optional parameters for water quality monitoring recommended on site specific basis are listed in Table 6.1. The justification for these parameters is outlined in Appendix 6.1. The measurement of some effluent parameters such as conductivity, sulphate or chloride concentrations may be useful as tracers to determine the extent of effluent mixing in the exposure area.

Calculation of Loadings

The MMER requires mines to record the monthly mass loadings of the MMER prescribed deleterious substances (MMER section 20). As part of effluent characterization for EEM, it is also recommended that mines calculate effluent loadings of the other parameters monitored. Loading can be calculated using the same method drescribed in the Section 20 of MMER by multiplying the mean effluent concentration of the parameter by the total volume of effluent discharge over the time period of interest (typically one year for effluent characterization).

6.1.4 Sampling Locations

Samples for effluent characterization will be collected from each final discharge point, identified by the owner or operator of the mine and in accordance with the MMER. "Effluent", as defined by the MMER, means mine water effluent, milling facility effluent, tailings impoundment area effluent, treatment pond effluent, treatment facility effluent - other than effluent from a sewage treatment facility - seepage and surface drainage that contains a deleterious substance.
Guidance for determining the sampling location(s) for effluent characterization is provided in the "Guidance document for the sampling and analysis of metal mining effluents" (EPS 2/MM/5, Environment Canada, 2001). This document focuses primarily on methods for collection of effluent samples from point sources (end of the pipe). If samples are to be collected from non-point sources, proposed sample collection locations and methods should be discussed with the Authorization Officer.

Water samples for water quality monitoring will be collected from at least one representative sampling station within the exposure area and at least one representative sampling station in the reference area four times per calendar year. These sampling stations will NOT likely be the same sampling stations used for biological monitoring. Sampling stations within the exposure should be located in an area surrounding the point of entry of effluent into the water from each final discharge. In selecting sampling stations, the owner or operator of a mine should take into consideration the location in the exposure area where effluent concentrations are the highest.

It is recommended that at least two to three water samples be collected at each sampling station to get an estimation of the variability and determine if concentrations of the contaminants are homogeneous within the sampling station. However, this may not be sufficiently robust to assess data statistically. More sampling stations within each area may help to better understand contaminant concentrations in the exposure area. At the minimum, a composite sample, consisting of few sub-samples spaced within the station, should be collected.

In addition to the above, the owner or operator of a mine will collect, concurrently with each fish and benthic invertebrate community study, no less than one water sample for water quality monitoring in each biological monitoring sampling area, and will record the measurements of parameters for water quality monitoring, as outlined in Table 6.1.

Water samples will be collected within each fish and benthic invertebrate sampling area at a representative station. It is strongly recommended, where the benthic and/or fish sampling areas are not in close proximity to the sampling station for water quality monitoring, that samples be collected concurrently at the sampling stations for the routine water quality monitoring. This will help to interpret the results of analyses of water samples collected in the benthic and/or fish sampling areas in comparison with the results of water quality monitoring.

6.1.5 Reporting

The results of effluent characterization and water quality monitoring will be submitted to the Authorization Officer as part of an Effluent and Water Quality Monitoring Report. This report will be submitted no later than March 31 of each year and will include, for the previous calendar year, the following:

- the dates when all samples were collected;
- the locations of the final effluent discharge points from which samples were collected for effluent characterization;
- location of the effluent discharge point used for sublethal toxicity testing;
- the latitude and longitude of sampling areas for water quality monitoring, in degrees, minutes and seconds and written descriptions that are sufficient to identify the sampling area;
- the results of effluent characterization;
- the results of sublethal toxicity testing (Chapter 7);
- the results of water quality monitoring;
- the methodologies used to conduct effluent characterization and water quality monitoring and the related method detection limits for analyses; and,
- a description of quality assurance and quality control measures that were implemented and the identification of any results that may interfere with the reliability of the data reported.

Since effluent samples for effluent characterization will be aliquots of samples collected for effluent monitoring, the measurements of pH, and the concentrations of the deleterious substances (arsenic, copper, total cyanide, lead, nickel, zinc, radium 226, and total suspended solids) will be available as part of the effluent monitoring reports from each mine submitted under section 21 of the MMER.

It is recommended that the Effluent and Water Monitoring Report include a comparison of water quality data in the exposure and reference areas. This comparison would examine all parameters included as part of water quality monitoring. In particular, this comparison should identify any parameters for which there are differences in measurements taken in the exposure and reference areas of more than a factor of two. When there are differences in measurements taken in the exposure and reference areas of more than a factor of two, it is recommended that the mine estimate and report the geographic extent for which this condition exists, based on expanded water quality monitoring or modelling. This is further discussed in Section 6.3.5

6.2 Effluent Characterization

6.2.1 Sampling Methods and Laboratory Analysis

Since effluent samples for effluent characterization will be aliquots of samples collected for effluent monitoring as part of the MMER, the sampling and chemical analysis considerations and recommended procedures provided in the **"Guidance document for the sampling and analysis of metal mining effluent"(EPS 2/MM/5, Environment Canada, 2001)** applies also to effluent characterization conducted as part of the EEM program. The volume of sample taken should be sufficient to allow for all required analyses and tests plus associated quality control samples (e.g. field duplicates, laboratory replicates and spiked sample). It is recommended that all samples be stored at 4°C and not frozen at any time before analysis (R. Turle, personal communication).

6.2.2 Methods for Determination of Thiosalts

Thiosalts are soluble sulphur/oxygen ions, which form as a result of the incomplete oxidation of sulphide minerals. They have the potential to be generated whenever sulphide minerals are in contact with oxygen, but in practice they tend to be formed during the processing of ores bearing sulphide minerals. If thiosalts occur in effluent, then once the effluent is discharged, the oxidation of the thiosalts is completed, resulting in the generation of sulphuric acid, and lowering the pH in the exposure area. Few mine sites in Canada have known thiosalt problems, but the potential for thiosalt generation may exist at many mine sites. As such, total thiosalt determination is optional for effluent characterization and water quality monitoring in the EEM program.

Methods have been developed for the determination of total thiosalts using a titration with sodium hydroxide, after mercuric chloride has been added to the test sample. Complete documentation of the method is available from Natural Resources Canada (Graham 1997).

The main concern with this method is that the samples should be analysed within 24 hours. As a result, total thiosalt analyses should be done on site. These analytical capabilities are likely restricted to sites with known thiosalts problems. Alternatively, samples may be frozen immediately after collection and be analysed within seven days. Longer storage time of frozen samples may affect thiosalts stability.

6.3 Water Quality Monitoring

The following guidance addresses specifically recommended procedures and methods for water quality monitoring in receiving environments.

6.3.1 Preparation for the field

The reagents for cleaning, operating or calibrating equipment, collecting, preserving and/or processing samples should be handled by appropriately qualified personnel and the appropriate data for health and safety (e.g., Material Safety Data Sheets) should be available.

Written protocols and standard operating procedures (including QA/QC requirements) should be readily accessible at all times, to ensure proper and safe operation of equipment. Data forms and log books should be prepared in advance so that field notes and data can be quickly and efficiently recorded. Extra forms should be available in the event of a mishap or loss. These forms and books should be water proof and tear resistant. Under certain circumstances, audio or audio/video recordings may prove valuable.

All equipment used to collect and handle samples should be cleaned and all parts examined to ensure proper functioning (e.g., on-site assembly or operation) prior to going into the field. A repair kit should accompany each major piece of equipment in case of equipment failure or loss of removable parts. Back-up equipment and sampling gear should be available.

Storage, transport, and sample containers, including extra containers in the event of loss or breakage, should be pre-cleaned and labelled appropriately (i.e., with a water-proof adhesive label to which the appropriate data can be added, with an indelible ink pen capable of writing on wet surfaces). The containers should have lids that are fastened securely. A sample-inventory log and a sample-tracking log should be prepared in advance of sampling. The responsibility for these logs should be assigned to one individual who will be required to monitor the samples from the time they are collected until they are analysed and disposed of, or archived.

6.3.2 Field Measurement of Water Quality Parameters

Standard *in situ* water quality parameters are dissolved oxygen, pH, conductivity, water temperature, and salinity. Total water depth at the water sampling station and water depth from which the water sample was collected should be recorded. Optical depth or transparency should also be measured in the field. Measurements of standard water quality parameters can be taken in the water directly, from a sample container in the boat, or on shore immediately after collection of the water, as long as the water is collected at the appropriate depth.

In shallow water bodies ≤ 2m deep, standard water quality parameters need only be measured at mid-depth. If the depth ranges from 2 to 4 m standard water quality measurements should be taken at two depth intervals, approximately 25 cm above the bottom and 25 cm below the surface. In deeper bodies of water, measurement of standard water quality parameters should be taken throughout the water column. Information on bottom depth and water column profiles of conductivity, pH, temperature, and dissolved oxygen should be obtained along one to five metre intervals (depending on total depth). For example, at a depth of 5 m measurements should be taken every metre, whereas at a depth of 25 m, measurements should be recorded at 5 m intervals.

Profiles can be facilitated through the use of a data logger (or equivalent) equipped with a dissolved oxygen probe and associated stirrer, as well as pH, conductivity, depth and temperature probes, which evaluate water column quality simultaneously. Such a unit is particularly useful for deeper evaluations (> 50 m). During profiling, the operator is able to visually review incoming data, noting particular areas of interest during descent and ascent of the unit (e.g., conductivity spikes, thermoclines, unusual data records, etc.). This information is recorded either manually or directly stored in the data logger. To supplement computer records, parameter readings should be stored manually onto field data sheets (every 2 or 5 m) depending on total depth profiled.

At shallow depths, hand held meters are often the most convenient way to measure *in situ* water quality parameters. They are light and several models are now available that can measure standard water quality parameters. The probes, and the cables connecting them to the hand held unit, can range from 2 to 5 m, limiting the use of such a unit. These meters tend to require more regular maintenance and calibration, therefore extra care should be taken to make sure that the meters are in proper functioning order, and calibration and maintenance logs should be kept on file.

Water depth can be measured indirectly using a sonar based fish finder, or directly using a calibrated tape, sounding cable, or rod. Recommended accuracy is as follows:

Water depth less than 2m: recommended accuracy of ± 25 cm
Water depth of 2 to 10 m: recommended accuracy of ± 50 cm
Water depth greater than 10m: recommended accuracy of ± 1 m

Optical depth is a measure of the transparency of water, and can be measured with a turbidity meter, in the field or in the laboratory. Optical depth can also be measured using a Secchi disk. The disk is 20 cm in diameter, and is painted white in two opposite quarters, and black in the other two. The disk is attached to a calibrated tape. To measure optical depth, the disk is lowered into the water in the shade until it has disappeared. It is then raised slowly, and the water depth at which it reappears is recorded. At least two measurements should be made at each station, and optical depth estimated based on the median value of the measurements. Measurements should be made at midday, and sunglasses should not be worn while measurements are made (Nielsen and Johnson 1983).

Water quality data should be screened on site during sample collection to ensure against the measurement and recording of false readings. This will permit the use of alternative instrumentation or instrument checks in the event of equipment or sampling error. All sampling and monitoring equipment should be checked and calibrated daily if necessary to ensure good working condition.

6.3.3 Collection of Water Samples for Laboratory Analyses

Water samples collected in the field and sent to a laboratory for analysis make up the bulk of the water quality monitoring program and involve the analysis of metals, nutrients, major anions and cations and several other general water quality parameters.

An ongoing debate regarding water quality monitoring is the issue of measuring total metal or dissolved metals. The AETE program examined this issue with the aim of determining if:

- the differences between metal concentrations in exposure and reference areas are more distinct using "total" rather than "dissolved" metals; and
- the linkage between water chemistry and in-stream biological effects was influenced by the choice of either "total" versus "dissolved" metals.

AETE concluded that, for both questions, there was no difference between measuring total or dissolved metals at the sites included in the AETE field program. AETE also highlighted the difficulties in filtering samples for dissolved metals analysis in the field (in a clean indoor work space), due to the risk of contamination during sample handling. As a result, AETE (ESG 1999) recommended the measurement of total metal concentrations for routine monitoring purposes. Based on this recommendation, ***total*** metals analysis is required during water quality monitoring.

In general, samples should be collected at two depth intervals: the subsurface (epilimnion) and near bottom (hypolimnion). If the water depth is ≤ 2m it is sufficient to collect water samples only at mid-depth or at least 15 cm below the surface. Samples collected below the surface of the water can be collected by hand directly into the sample bottle.

Water collections at discrete depths should be facilitated through the use of non-metallic 2-16L Van Dorn or 0.5-8L Kemmerer samplers (or other type, e.g. Niskin sampler). For streams, depth integrated samplers which are representative of the suspended sediment and related substances can be used. These samplers can be used from a boat, bridge or ice surface, and usually require two persons for safe operation. For very deep samples, a peristaltic sampler could be used in preference to other types. If other types of samplers are chosen, one criteria for consideration is that they should be Teflon-coated.

The water sampler should be triple rinsed with the water from the sampling station between each sample. In addition, it is recommended that sampling in the reference area be completed first to avoid any potential contamination of the sampler with water from exposure area. The sampler should be double rinsed with reagent grade weak nitric acid between sampling areas, particularly if it is not possible to complete sampling in the reference area first.

When collecting water samples, it is important to use as many of the following ultra trace techniques and proper water sampling protocols as possible:

- sampling should proceed from the least contaminated to the most contaminated station
- sample bottles and caps should be rinsed three times prior to water collection
- no preservative should be placed into the sampling bottles prior to sample collection;
- samples should be collected with the bottle mouth facing up-current away from the sampler's hand;
- sample collectors should wear unlined latex or nitryl gloves to avoid contamination of the sample
- caps of water containers should be held lid-down during sample collection.

6.3.4 Sampling Handling, Storage and Analysis for Water Quality Monitoring

6.3.4.1 Sample Handling and Preservation

Table 6.2 lists the handling, container and preservative recommendations for specific parameters. Note that to reduce the number of samples collected, several analytes can be analysed from one sample bottle. Prior to sample collection the list of parameters should be discussed with the laboratory to determine the number and type of sample bottles required. When collecting samples, it is useful to have a checklist similar to Table 6.3, which lists the collection bottles, corresponding analytes, and whether or not a preservative is required; as a sample is collected it should be checked off the list. In certain situations, a maximum holding time of seven to ten days (major cations and anions, nitrate/nitrite, dissolved organic carbon) may be problematic. If the shipping of a mine's water samples has been unavoidably delayed but sample integrity was retained, the Authorization Officer should be notified without delay.

Sample Bottle Types

Except where noted in Table 6.2, HDPE bottles are highly recommended for samples for chemical analysis for water quality monitoring. Hall (1998) concluded that bottles made of high density polyethylene (HDPE) showed the least amount of contamination, and are also the least expensive. These bottles are also less likely to break than glass. New HDPE bottles can be used without rigorous cleaning (only rinsing with de-ionised water); however, a rinse with weak HNO_3 is recommended (Hall 1998). Any reputable laboratory should have these containers available.

Sample Preservation

Some samples require preservation to ensure stability of target compounds during transportation and storage or to eliminate substances which may interfere with the analysis. In some cases, preservation of the sample is optional, and if selected, will allow for a longer storage period before analysis should be initiated. Generally, samples requiring preservation should be preserved immediately upon collection. Note however, that field preservation of samples can be a significant source of contamination, and care should be taken during preservation to ensure sample integrity.

Where samples are to be preserved to a fixed set-point (pH, colour) care should be taken to ensure that the set point has been reached by using detection techniques such as confined range pH paper, pocket/portable pH meters, standard colour comparison charts, etc. The use of these techniques and/or devices should not contaminate the sample.

Preservation methods were evaluated by Hall (1998) and it was found that acidification to 0.4% HNO_3 (or $pH < 2$) should maintain most metals in water samples for at least one month at room temperature. Preservatives should be added in quantities that do not dilute the sample, generally at less than 1% of the total sample volume. Use only highly pure reagent grade chemicals or better as preservatives in order to avoid contamination. Usually the rate of addition of concentrated nitric acid is 0.1-0.2mL/100mL of sample.

Bottles should **not** be pre-charged with preservatives as concentrated preservatives can cause leaching of metals from the plastic (Hall 1998). This should be specified when bottles are ordered, since some laboratories provide pre-charged bottles.

Water samples for mercury analysis should be collected and preserved separately. The best preservation of Hg was identified as 0.5% BrCl, while 2% HCl or 0.04% $K_2Cr_2O_7$ were considered questionable (Hall 1998).

CAUTION: Acid preservation of samples suspected of containing cyanide or sulphide MUST BE carried out in a well ventilated area.

All sample containers and required preservatives should be provided by the laboratory hired to conduct the analyses of samples. Bottles should preferably be new and unused and should be purchased as certified clean. If bottles are reused, they should be cleaned by a documented cleaning procedure with a bottle lot number control system, and cleanliness should be demonstrated by the use of blanks.

Table 6-2: Water sample handling recommendations regarding containers, preservatives and holding times for specific water quality parameters

Parameter	Container Specifications/Suggested Sample Size	Preservative (in addition to cold storage)	Max. Holding Times[3]
Total Metals	500 mL HDPE[1]	5 mL 20% nitric acid	6 months
Dissolved metals	500 mL HDPE	5 mL 20% nitric acid after filtration	6 months
Total mercury	125 mL glass	0.5% BrCl	28 days
Major cations and anions (including fluoride)	1 L HDPE	None	7 days
Total Suspend Solids	1 L HDPE	None	14 days
Total Organic Carbon	250 mL HDPE	H_2SO_4, pH<2	30 days
Dissolved Organic Carbon	100 mL amber glass, TFE[2]-lined cap or HDPE	Filter, 1 mL 1:1 sulphuric acid to pH 2	10 days
Conductivity	1L HDPE	None (ideally measured *in situ*)	28 days
PH	1 L HDPE	None (ideally measured *in situ*)	14 days
Hardness	1L HDPE	None	6 months
Alkalinity	1L HDPE	None	14 days
Acidity	1L HDPE	None	14 days
Dissolved oxygen	125 mL HDPE	None, fill to top	Measure in field if possible
Turbidity	1 L HDPE	None	48 hr
Ammonia-nitrogen	500 mL HDPE or glass	2 mL 1:1 sulphuric acid, pH 2	30 days
Nitrate nitrogen	1 L HDPE or glass	None	7 days
Nitrite nitrogen	1 L HDPE or glass	None	7 days
Total cyanide	1 L HDPE	10 mL 6N NaOH	14 days
Total phosphorus	500 mL HDPE or glass	2 mL 1:1 sulphuric acid, pH 2	30 days
Radium 226	250 mL HDPE	HNO_3, pH<2	30 days
Thiosalts	1 L HDPE	Freezing	7 days

1 HDPE – high density polyethylene

2 TFE - tetrafluroethylene

3 Maximum time chosen from either EPA, MISA and Contaminated Sites methodologies.

Table 6-3: Typical checklist of bottle requirements for parameters for water quality monitoring.

Bottle #	Bottle type	Preservative	Parameters
1	125 mL HDPE	HNO_3, pH<2	Total metals
2	125 mL glass	HNO_3 & BrCl	Mercury
3	1L HDPE	None	General chemistry: Cl, F, SO_4, NO_3, major ions, thiosalts, TSS, DOC, hardness, pH, conductivity, alkalinity, acidity
4	125 mL HDPE	NaOH, pH>12	Cyanide
5	250 mL HDPE	H_2SO_4, pH<2	Ammonia, TOC, total phosphorous
6	250 mL HDPE	HNO_3, pH<2	Radium 226

6.3.4.2 Sample Shipping and Storage

Samples should be transported to a laboratory as soon as possible after collection. Immediately after collection samples should be stored in a cooler containing well-sealed ice packs. It is also recommended that samples be shipped at 4°C to minimise degradation of analytes. Shipping coolers should be stocked with ice packs to ensure that samples are kept cold. Speed in shipping by pre-arranged transport is recommended. These logistical questions should be addressed prior to collecting the samples, particularly at remote locations.

Analyses should be completed within the accepted storage times (Table 6.2). Storage time is defined as the time interval between the end of the sample collection period (typically 24 hours for composite samples) and the initiation of analysis. All samples should be stored for as short a time interval as possible and under conditions that minimise sample degradation. Samples should be maintained at temperatures **above their freezing point** and under 10°C, with minimal exposure to light. Samples digested for metals analysis may be maintained in a sealed container and analysed within 30 days.

6.3.4.3 Unique Sampling Requirements for Some Chemical Groups

Cyanide: Samples containing strong oxidising agents (e.g. chlorine) should be neutralised as soon as possible after sample collection to prevent oxidation/degradation.

Ammonia: Samples containing strong oxidising agents (e.g. chlorine) should be neutralised as soon as possible after sample collection to prevent oxidation/degradation.

Elemental Scan: Changes in sampling techniques should be kept to a minimum: once a valid mode of sampling has been selected, it should be maintained and changed only if necessary.

6.3.4.4 Laboratory Analyses of Samples

Laboratory analyses should be carried out in a qualified laboratory by trained personnel operating under quality controlled conditions and using documented standard operating procedures (SOPs). It is recommended that laboratories being used to generate data for the mining EEM program are accredited by the Standard Council of Canada through the Canadian Association for Environmental Analytical Laboratories (CAEAL) or the Ministère de l'Environnement (MEQ) du Québec. These accreditation organisations follow International Standards Organisation (ISO/IEC) standards, which focus on the quality system and practices of a laboratory.

Tables 6.4 and 6.5 list commonly accepted analytical methods and instrumental or method detection limits where available. It is recognized that obtaining these detection limits may be difficult in some situations, but it is recommended mines target the limits as outlined. These detection limits will be re-examined after the first three years of the EEM program to assess their appropriateness in the mining environment.

For the analysis of metals, Inductively Coupled Plasma (ICP) methods combined with either a mass spectrometer (MS), atomic emission spectrometer (AES) or optical emission spectrometer (OES) as a detector are all relatively common. Of these methods ICP-MS is the most ideal for water analysis due to its sensitivity, simple spectra and ability to obtain isotopic information and scan multiple elements. Graphite furnace atomic absorption spectrometry (GFAAS) is also used to measure metals in water and is sensitive. However, this method is disadvantageous in that it is essentially a single-element technique and interference protocols have to be strictly adhered to. If a larger group or suite of metals is being analysed, ICP-MS is usually a more cost-effective method.

Note that, since concentrations of parameters tend to be higher in effluent than in the exposure area, it may be possible to use analytical methods for effluent characterization that have higher detection limits.

Other technologies for analysing metals in water include Anodic Stripping Voltametry (ASV), Ion Selective Electrodes (ISE), Ion Exchange Resins (IER) and Ion Chromatography (IC). These methods were reviewed as part of the AETE program and it was concluded that ICP and AAS (Atomic Adsorption Spectrometry) techniques should continue for metals analysis until some of the alternatives are further developed and validated (ESG 1999).

Note however, that ion chromatography can be used for analysis of major ions, specifically when a suite of anions is analysed together, the ion chromatography presents advantages over colorimetric methods.

The analytical methods selected should meet the criteria in this document plus any other objectives identified by the mine (or those acting on the mine's behalf) or Environment Canada. The project manager and the laboratory need to confirm what parameters of interest will be measured and that holding times can be met. The laboratory and analysis methods should be selected and discussed before the sample is collected to ensure that the laboratory sample requirements are met.

Analytical methods selected should be generally accepted and in common use in laboratories in Canada. The overall method principle should be peer reviewed and widely published so that it can be located easily for details.

Table 6-4: Detection limits for methods used to determine metal concentrations in water samples

Parameter (Total metals)	**ICP/MS[1] APHA Method 3125 (IDL[2] µg/L)**	**Flame-AAS[3] AHPA Method 3111, except where noted (IDL µg/L)**	**GFAAS[4] APHA Method 3113 (EDL[5] µg/L)**	**ICP-OES[6] NLET[7] Std. Method 2-2001, except where noted (MDL[8] mg/L)**	**ICP/MS Series 200.8 (IDL µg/L)**
Aluminium	0.03	0.1	3	0.002	0.05
Arsenic	0.025		1	0.0001 (2-2200 ICP-AES[9])	0.9
Cadmium	0.003-0.006	0.002	0.1	0.0001	0.1
Copper	0.003-0.004	0.01	1	0.0002	0.03
Iron	No ICP/MS method	0.02	1	0.0004	No ICP/MS method
Lead	0.005	0.05	1	0.005	0.08
Mercury		(APHA 3112 Cold-vapour AA[10])		(2-2300 Cold vapour AA)	0.2
Manganese	0.002	0.01	0.2	0.002	0.1
Molybdenum	0.003	0.1	1	0.003	0.1
Nickel	0.004-0.025	0.02	1	0.004-0.025	0.2
Selenium	0.064-0.093		2	0.064-0.093 (may also use 2-2200 ICP-AES	5
Uranium	0.001-0.032			0.001-0.032	0.02
Zinc	0.017-0.020	0.005		0.017-0.020	0.02

1. ICP/MS – Inductively coupled mass spectrophotometry
2. IDL – instrumental detection limit. Determined by instrument manufacturer. Actual MDL may be different.
3. Flame AAS – Flame atomic absorption spectrometry
4. GFAAS – Graphite furnace atomic absorption spectrometry
5. EDL – Estimated detection limit
6. ICP-OES – Inductively coupled
7. NLET – National Laboratory for Environmental Testing
8. MDL – Method detection limit
9. ICP-AES – Inductively coupled optical emission spectrometry
10. Cold vapour AA – Cold vapour atomic absorption spectrometry

Table 6-5: Methodologies and detection limits used to determine concentrations of water quality parameters .

Other Parameters:	NLET Method (MDL)	Description	APHA Std. Method (MDL)	Description
Calcium	1-1050 (0.02-0.1mg/L)	Atomic absorption spectroscopy	3030/3120 (0.005-1 mg/L)	Acid digestion followed by ICP emission spectroscopy
Chloride	1-1070 (0.05 mg/L)	Colorimetric, ferric thiocyanate	4110 B (0.1 mg/L)	Ion chromatography with chemical suppression of eluent conductivity
Fluoride	1-1120 (0.01 mg/L)	Ion-selective electrode	4500-F-C (0.05 mg/L)	Ion-selective electrode method
Magnesium	1-1140 (0.01-0.1 mg/L)	Atomic absorption spectroscopy	3030/3120 (0.005-1 mg/L)	Acid digestion followed by ICP emission spectroscopy
Potassium	1-1220 (0.02 mg/L)	Flame emission photometry	3030/3120 (0.005-1 mg/L)	Acid digestion followed by ICP emission spectroscopy
Sodium	1-1240 (0.02 mg/L)	Flame emission photometry	3030/3120 (0.005-1 mg/L)	Acid digestion followed by ICP emission spectroscopy
Sulphate	1-1250 (0.2 mg/L)	Colorimetric, methyl thymol blue	4500-SO_4-E (0.5 mg/L)	Turbidimetric method
Total cyanide			EPA 335.3 (0.001 mg/L)	Automated continuous flow digestion and distillation followed by colorimetry
Ammonia-nitrogen	1-1161 (0.001 mg/L)	Colorimetric, phenate	4500-NH_3-F (0.005 mg/L)	Phenate method:Some water samples will require distillation. Manual measurement of indophenol blue at 630 nm.
Nitrate-nitrogen	1-1182 (0.001 mg/L)	Cadmium reduction	APHA-NO_2-B (0.006 mg/L)	Colorimetric method, nitrite is determined through addition of reagents to form an azo dye, meaurement of color at 520 nm.
Total suspended solids (TSS)			APHA 2540 D (10 mg/L)	Filter sample through washed, dried, pre-weighed Whatman GFC glass fibre filter and dry filter solids at 105° to dryness. Determine weight change in filter gravimetrically.

* Control limit (± 3 SD) based on recovery of surrogate compound spikes.

Table 6.5: Methodologies and detection limits used to determine concentrations of water quality parameters in water (continued).

Other Parameters:	NLET Method (MDL)	Description	APHA Std. Method (MDL)	Description
Total organic carbon (TOC			APHA 5310 (1 mg/L)	Prep: Acidify to pH 2 and purge with nitrogen to remove inorganic carbon. High temperature combustion with infrared detection of CO_2
Dissolved organic carbon (DOC)			APHA 5310 B (1mg/L)	Prep: Acidify to pH 2 and purge with nitrogen to remove inorganic carbon. Pass through 0.45 µm pore diameter filter and analyse filtrate. High temperature combustion with infrared detection of CO_2
Total phosphorus	1-1190 (0.0002 mg/L)	Colorimetric, stannous chloride	APHA 4500-P-B-E (0.002 mg/L)	Sulphuric acid-nitric acid or persulphate digestion followed by ascorbic acid method
Total hardness	1-1130 (not applicable)	Calculated	APHA 2340 B (NA)	Calculation method from Ca and Mg data (i.e. cations)
pH	1-1600 (NA*)	Electrometric	APHA 4500-H, B (NA)	Electrometric method
Conductivity	1-1800 (0.2 µS/cm)	Wheatstone bridge conductance cell	APHA 2510 B (0.2 µS/cm)	Laboratory method using conductivity meter
Alkalinity	1-1030 (0.1 mg/L)	Conductometric titration	APHA 2320 B (5 mg/L)	Titration method
Acidity	1-1001 (0.1 mg/L)	Electrometric titration	APHA 2310 B (5 mg/L)	Titration method
Radium 226[1]	0.01 Bq/L	Spectometr, radon emanation , α counting		co-precipitation with $BaSO_4$

* Control limit (± 3 SD) based on recovery of surrogate compound spikes.

1. Methodologies and detection limit recommended for effluent monitoring: to be verified for water quality monitoring

6.3.5 Comparison of Water Quality Data in Exposure and Reference Areas

6.3.5.1 Comparison of Results

It is recommended that the Effluent and Water Monitoring Report include a comparison of water quality data in the exposure and reference areas. This comparison would examine all parameters included as part of water quality monitoring. In particular, this comparison should identify any parameters for which there are differences in measurements taken in the exposure and reference areas of more than a factor of two.

The factor of two for exceedance of concentrations from the reference area is intended to ensure that differences between exposed and reference area concentrations are real differences, and not just differences that may be attributed to such factors as: low concentrations of target contaminants, analytical variability, small minimal sample size (n = 4) and seasonable variability. At sites where the reference area is on a different water body or watershed than the exposed area, the factor of 2 difference may not be applicable.

Determination of whether or not concentrations are different between the exposure and reference areas should be based on the median value of a minimum of four samples collected over a 12 month period from the same exposure and reference area locations. The median in a set of n measurements $y_1, y_2, y_3, \ldots y_n$ is defined to be the value of y that falls in the middle when the measurements are arranged in order of magnitude. If there is an even number of measurements, then the median is the value of y halfway between the two middle measurements. If larger data sets ($n >> 4$) are available, then determination of whether of not concentrations are elevated in the exposure area could be based on a statistical test such as mean or median greater than the 95% confidence interval or greater than 2 standard deviations. If there is adequate pre-mining water quality data in the exposure area, then pre-mining data may be used as a basis for comparison.

In cases where there are differences of more than a factor of two, it is recommended that the mine estimate and report the geographical extent for which this condition exists, based on expanded water quality monitoring or modelling. However, before completing such an estimate, there are a number of factors which should be considered:

1) *Site specific water quality objectives:* if there is a site specific water quality objective for a particular parameter, and that objective is exceeded in the exposure area, then the extent of this exceedence should be determined, regardless of the concentration in the reference area.

2) *Water quality guidelines:* if there are water quality guidelines for a particular parameter (e.g., federal or provincial), and the concentrations of that parameter in the exposure area are more than a factor of two greater than concentrations in the

reference area, and are greater than thewater quality guideline, then the extent of this exceedence should be determined.

The Canadian Council of Ministers of the Environment (CCME) water quality guidelines for parameters for water quality monitoring are listed in Table 6.6. Several provinces have also developed water quality guidelines, and in cases where CCME and provincial water quality guidelines exist for a particular parameter, the provincial guidelines take precedence, although both should be reported.

3) *Detection limits:* In cases where the generic water quality guideline for a parameter is close to the analytical method detection limit, and concentrations of the parameter in the study area are close to the guideline, a factor of 2 difference may not be meaningful (as a result of analytical uncertainty close to the detection limit). In such cases the Authorization Officer should be consulted. McQuaker (1999) provides a comparison of achievable detection limits with generic water quality guidelines and, for most parameters, method detection limits significantly less than the water quality guideline (at least a ratio of 1:10 MDL: WQG) are available. However, McQuaker concluded that there are some parameters (arsenic, cadmium, mercury, selenium, silver and cyanide) for which a MDL at least ten times lower than the WQG is not currently achievable and may not be commercially available. As the MDL/WQG ratio decreases, the measurement uncertainty increases; beyond a ratio of 1:2, the results are not considered to be statistically significant.

4) *pH:* For pH, a difference of a factor of 2 would not be significant, since the pH scale is logarithmic. If there is a site-specific objective for pH, and if the pH in the exposure area is outside the range specified in the site-specific objective, then the geographical extent of pH values outside the range of the site-specific objective should be determined. If there is no site-specific objective for pH, and if the pH in the exposure area is more than 0.5 pH units different than the pH in the reference area, and is also outside water quality guidelines (e.g. CCME 6.5 to 9.0), then the geographical extent of the exposure area within which pH is more than 0.5 pH units different than in the reference area should be determined.

5) *Location of the reference area:* At sites where the reference area is on a different water body or watershed than the exposure area, the factor of 2 difference may not be applicable. If it is felt that this is the case, the Authorization Officer should be consulted.

6.3.5.2 Estimation of Extent of Elevated Concentrations

Two methods may be used to estimate the extent of elevated concentrations:
- direct measurement
- modelling

Table 6-6: Summary of CCME water quality guidelines for freshwater aquatic life for key parameters monitored as part of the EEM program.

Key Parameters	CCME Water Quality Guideline (µg/L)	Key Parameters	CCME Water Quality Guideline (µg/L)
arsenic	5.0	aluminium	5 at $pH < 6.5$; $[CaCO_3] < 4$ mg/L; $DOC < 2.0$ mg/L 100 at $pH \geq 6.5$; $[Ca^{2+}] \geq 4.0$ mg/L; $DOC \geq 2.0$ mg/L
copper	2 at H[1] = 0 - 120 mg/L (1) 3 at H = 120 - 180 mg/L 4 at H > 180 mg/L	cadmium	0.017[2] at H = 48.5 mg/L
lead	1 at H = 0 - 60 mg/L 2 at H = 60 - 120 mg/L 4 at H = 120 - 180 mg/L 7 at H > 180 mg/L	iron	300
		mercury	0.1
nickel	25 at H = 0 - 60 mg/L 65 at H = 60 - 120 mg/L 110 at H = 120 - 180 mg/L 150 at H > 180 mg/L	molybdenum	73
		ammonia-nitrogen	guideline is for total ammonia 1370 at pH 8.0; 10°C 2200 at pH 6.5; 10°C
Zinc	30		
Radium 226	no CCME guideline	nitrite	60 (nitrite)
Total cyanide	5 (as free CN)	manganese	no CCME guideline
pH	6.5 to 9.0	selenium	1
Total Suspended Solids	guidelines based on increase from baseline.	Uranium	no CCME guideline
		fluoride	no CCME guideline
		total phosphorus	no CCME guideline

[1] H = hardness as CaCO3

[2] cadmium guideline is hardness dependant, according to the following formula: guideline = $10^{\{0.86[\log(hardness] - 3.2\}}$

1) *Direct Measurement*

Direct measurement requires an increase in the number of sampling stations within the exposure area to determine where within the exposure area concentrations of contaminant(s) of concern are no longer elevated. The number of additional stations needed would be determined on a site-specific basis, but generally, a minimum of three stations would be needed:

- the exposure area station used for water quality monitoring

- 2 or more further stations that bracket the location within the exposure area where the concentrations of parameter(s) of concern are no longer expected to be elevated.

2) Modelling

If the seasonal variations of concentrations of the key parameter of concern in effluent and the exposure area are well understood, and if seasonal variations in effluent and receiving environment flow are well understood, then it may be possible to predict the location within the exposure area where the concentrations of the parameter(s) of concern are no longer expected to be elevated.

6.4 Quality Assurance and Quality Control for Water Quality Monitoring

Quality assurance and quality control (QA/QC) is a documented system incorporating adequate review, audit and internal quality control. The objective of a QA/QC program is to ensure that all field sampling and laboratory analyses produce technically sound and scientifically defensible results.

Quality assurance is a planned system of operations and procedures whose purpose it is to provide assurance to the client that defined standards of quality are being met. Analytical quality assurance defines the way in which tasks are to be performed to ensure that data meet pre-defined data quality goals. These tasks include not only the analysis itself but all aspects of sample handling and data management.

Quality control is an internal aspect of quality assurance and it refers to activities that are being carried out to implement the quality assurance plan. It includes the techniques used to measure and assess data quality and the remedial actions to be taken when data quality objectives are not realised. Quality control activities ensure that the quality of results meet the needs of the client. Specific quality control measures include mechanisms such as setting data quality objectives and standard operating procedures that apply to field and laboratory aspects of the study.

6.4.1 Standard Operating Procedures

Standard operating procedures (SOPs) are fundamental to any QA/QC program. All field and lab procedures should be conducted according to SOPs to ensure quality control. SOPs should describe in detail:

- the sampling methods and procedures, sample handling, labelling, equipment, preserving, record-keeping and shipping requirements of the field programs; and

- the analytical methods and procedures, sample handling, labelling, equipment, test system implementation, record-keeping, etc., of all laboratory analyses.

Each SOP should be a written method accessible to each analyst. SOPs should be based on procedures developed by a standard-setting organisation such as Environment Canada, the U.S. EPA, the American Society for Testing and Material (ASTM), or the American Public Health Association (APHA). Where methods are not well validated, it is recommended that the SOP be thoroughly referenced to the relevant literature and contain all the elements outlined in Canadian Association for Environmental Analytical Laboratories (CAEAL) Inc. (1991), where applicable. In-house validation data should be appended to the SOP and should contain the quality assurance and quality control procedures. These include the types and frequencies of quality control samples to be analysed, expected levels of precision, accuracy, recovery, and method detection limits.

While chemical analysis procedures tend to be reasonably well-documented, sampling procedures in general, and sampling design in particular, are often overlooked. Sampling error is usually a large component, and often the dominant component of uncertainty in environmental measurements. Standard operating procedures that include field operations will help to reduce this uncertainty or at least ensure that it is quantified.

Emphasis should be placed on measures to prevent inadvertent contamination of samples and to ensure sample integrity. In addition, SOPs will specify the proper preparation of all sampling gear and supplies, and the proper calibration of all instrumentation (such as meters).

6.4.2 General Aspects of Quality Control in the Field

General quality control aspects of a field sampling program are:

- all personnel involved in field procedures should have appropriate education and training;
- sampling methods should be consistently applied among sites throughout the study;
- samples should be collected according to SOPs that should be available to personnel at all times during the field study;
- sampling equipment should be appropriate for the habitat being studied, properly cleaned, and accompanied by the appropriate documentation (i.e., manual, calibration and maintenance schedule);
- all samples should be properly labelled as to date, location, type, number and collector;
- samples should be in the proper container with the appropriate preservative or fixative if necessary;
- field technicians should maintain detailed field notes using indelible ink and waterproof notebooks;
- personnel should use chain-of-custody/sample submission forms and custody seals for contaminant samples;

- personnel should follow appropriate shipping and storage methods; and
- standardised field collection forms should be used during the field program.

6.4.3 Field Aspects of Quality Assurance

Field quality assurance for water quality monitoring should be achieved through several differing and supporting methodologies which include duplicate readings, comparison of readings with known standards, collection of profile samples for analytical evaluation and parameter evaluation using alternate equipment (e.g., Hanna CTD meter, thermometer, etc.).

Some of the most common quality problems are the result of mislabelling or switching bottles, failure to add proper preservatives, improper storage conditions, sample contamination from sampling equipment and exceeding the holding time. Each sample should be clearly labelled in a manner that identifies the sample and distinguishes it from all other samples. Labels should be filled out in indelible ink and fixed to the sample container such that it will not fall off when wet or during transport.

The field log book is an integral part of the sampling program and forms the basis of the sampling report. Items documented in the log book are often highly relevant to the interpretation of the laboratory data. Any deviations from the sampling plan or any other observation about the sample or the sampling locations should also be noted in the log book. Some common deficiencies in field log books are failure to make planning notes, failure to make notes at the time events occur, failure to sign and date entries and failure to write legibly.

6.4.4 Quality Assurance During Sample Handling, Shipping and Storage

CAEAL (1991) recommends the following with respect to quality assurance during sample handling, shipping and storage.

a) ***Chain of Custody*** Chain-of-custody forms should be used in the transportation of samples, especially in cases where several contracted parties are involved in the sampling, shipping, and analysis of the samples. An example of a chain-of-custody form is presented in Figure 6.1.

b) ***Sample Inspection*** The condition of each sample should be noted upon receipt. Discrepancies between required sample conditions and the observed conditions should be recorded in a logbook or on a computer file. The samples should be preserved immediately if submitted without preservatives, and a record made of the preservation methodology. (It is preferable to preserve samples in the field.)

c) ***Sample Tracking*** Samples should be assigned a unique number or code to identify the sample in a tracking system. The sample tracking system should identify the sample, the source, the date of receipt, analyses, due date, and any other pertinent information. A computerised laboratory information management system (LIMS) is recommended for tracking samples in laboratories processing large numbers of samples for a variety of clients.

d) ***Sample Storage*** Samples should be stored in an assigned location in a refrigerator or sample storage area accessible only to authorised personnel. Samples should be refrigerated at 4°C±2°C, where applicable, and removed only for inspection, logging, and analysis. The temperature of the refrigerator should be measured and recorded daily.

Figure 6-1: Example of a chain-of-custody record (modified from U.S. EPA 1988)

Survey				Samplers (signature)					
Station Number	Station Location	Date	Time	Sample Type Water		Preservative added (initial)	Seq. No.	No. of Containers	Analysis Required
				Comp	Grab				

Relinquished by (signature)	Received by (signature)		Date/time
Relinquished by (signature)	Received by (signature)		Date/time
Relinquished by (signature)	Received by (signature)		Date/time
Relinquished by (signature)	Received by (signature)		Date/time
Dispatched by (signature)	Date/time	Received for laboratory by:	Date/time
Method of shipment:			
Distribution Orig. - accompany shipment 1 copy - survey co-ordinator field files			

6.4.5 Use of Blanks and Duplicate Samples

The use of blanks and duplicate samples in the field and laboratory is an important component in a quality control program.

Field blanks and field duplicates are essential throughout the execution of a field program involving the collection of water. Field QC samples are used to establish whether any errors are being introduced during the sampling process so that corrective action can be taken if necessary. Field QC samples are distinct from laboratory QC samples as they measure sampling effects rather than laboratory effects.

Field blanks are used to check contamination from all potential sources of contamination of the sample. These include possible contamination of sample bottles, caps, preservatives, equipment, filter paper (if samples are to be filtered), atmospheric contamination, sampling techniques, and analysis. Field blanks are collected by obtaining blank water (i.e. Milli-Q water) from the laboratory conducting the analyses, transporting the water to the field, and taking it through all sample collection, handling and processing steps that the test samples undergo (e.g., transfer to a sample container, preservation and exposure to the environment). Field blanks are transported, stored and analysed in the same manner as test samples (McQuaker 1999).

Duplicate samples should be taken to verify analytical results and equipment reliance. Field duplicates are used to evaluate homogeneity of the sample site and the ability of the sampling system to take the sample the same way every time. A field duplicate is a completely separate sample, not a split of a single sample into two bottles. Duplicate samples should be treated as blind samples, and are not identified to the laboratory.

The last type of QC sample is the trip blank, also referred to as travel or transport blanks. Trip blanks are used to check contamination from sample bottles, caps and preservatives during transport, storage and analysis. A sample bottle is filled in the laboratory with blank water (i.e. Milli-Q water) and preserved in the same manner as the test samples will be. Trip blanks are transported to the field with regular sample bottles and submitted to the laboratory unopened, together with the test samples. They are opened at the time of analysis, and analysed in the same manner as the samples (McQuaker 1999).

Field and trip blanks, and duplicate field samples should be collected at a frequency of 5 to 10% of the total number of samples. Therefore, if a total of 10 water quality stations were being sampled, only 1 of each of the QC samples should be required. This proportion can be increased if necessary, to monitor errors due to sampling and matrix homogeneity. If field and trip QC samples are not used, any inaccuracy introduced due to sampling will go undetected or be inappropriately attributed to the analytical laboratory. The use of blanks and duplicate samples in the laboratory are further discussed in Section 6.4.8. Table 6.7 summarises recommended use of blanks and duplicate samples in the field and the laboratory, for larger sampling programs. For routine sampling, with one station from the exposure area and one from the reference area, it is recommended that a single field blank be submitted together with the test samples. In such cases, these samples will be analysed by the laboratory as a batch, together with samples from other clients. The laboratory will achieve necessary internal QC using the complete batch.

Table 6-7: Summary of recommended use of blanks and duplicate samples in the field and laboratory. Numbers are based on a batch of 20 samples.

Parameter	Number of Samples	Internal or Field QC	Control limits	Description
Field blank	1	Field		Checks contamination as a result of sample handling. One per day per matrix
Trip blank	1	Field		Tests validity of sample preservation and storage conditions. One per day per matrix
Field Duplicate	1	Field		Used to evaluate homogeneity of the sample site and the ability of the sampling system to take the sample the same way every time.
Method Blank	1	Internal	< D.L. or < 0.1 of sample conc.	Checks contamination from reagents and procedures[a]
Laboratory Duplicate Sample	1	Internal		Checks precision of sampling process. One per day per matrix type[a]
Glassware Proof	1	Internal	< D.L. or < 0.1 of sample conc.	Checks contamination of lab glassware used during processing[a]
Standard reference material (SRM)	1	Internal		Checks accuracy of method[a]
Matrix spike	1	Internal	75 to 125%	Used interchangeably with SRM[b]
Calibration control:				
Within run (blank and mid-range standard	1	Internal	10% drift max.	Statistical control over calibration can be confirmed between runs by means of two control standards, A and B, and within-run by means of blanks and midrange standards. (King 1976)
Between runs (20% and 80% full scale)	2 per run	Internal	± 5% of target value	

[a] Intrinsic to every batch of 20 samples

[b] Used interchangeably with SRM if SRM is not available.

6.4.6 General Aspects of Quality Control in the Laboratory

The following lists the general quality control aspects of the laboratory analyses performed on water samples:

- data should be verified and validated through transcription checks; chemical data will be verified by reference to the analytical laboratory QA reports accompanying the data.
- data analyses will be repeatable and robust and will be cross-checked with data quality objectives.
- data analyses will be rigorous and defensible and will include the rationale for all statistical analyses and data transformations.

6.4.7 Details on Quality Control Aspects of Laboratory Analyses

Analytical quality control procedures are designed to demonstrate statistical control over calibration, precision, accuracy/bias, and recovery (CAEAL) 1991.

Statistical control over these parameters can be demonstrated by running specific quality control samples during each analytical run (see Table 6.7). The results of these QC samples are compared statistically with confidence intervals calculated from historical data. These confidence intervals or control limits are normally calculated at three standard deviations of the mean of the controlled variable. Warning limits are frequently set at two standard deviations. Indicators of a run considered out of control include the following:

- two successive results for method blanks, laboratory duplicates, standard reference materials, spiked blanks, calibration control samples, or organic surrogate recoveries
- one of these results outside the control limits.

Quality control data can be plotted on appropriate control charts. Control charts are graphic presentations of the QC data as a function of time or consecutive run number. Control charts demonstrate trends in time and provide graphic evidence of long-term statistical control of the analysis. Control limits and control charts are described in detail in ASTM (1986).

6.4.7.1 Good Laboratory Practices

Well established, good laboratory practices (GLPs) should be followed. The following is a brief listing of recommended laboratory practices. A description of GLPs can be found in greater detail in ELAP (1988).

- records on reagent preparation should be maintained in a log book. Prepared reagent containers should be labelled with the reagent, its date of preparation, expiry date, and the person responsible.
- instruments should be maintained or serviced on a regular basis. Maintenance records should be kept in a log book.
- written instructions should be available for all instruments.
- standard procedures for cleaning glassware and containers should be followed.
- routine checks of the purity of the distilled water should be conducted and documented. Distilled/de-ionised water should be checked on a conductivity meter at least daily.
- chemical reagents should meet the purity requirements of each analytical method.
- reagents and solvents should be stored according to the manufacturer's directions.
- working standards and stock solutions should be checked to determine changes in concentration.
- reagents should be prepared and standardised against primary reference standards.
- the temperatures of all refrigerators and incubators should be checked daily and temperature excursions should be recorded.
- each oven should have a dedicated thermometer and the temperature should be checked prior to and following each use.
- proper volumetric glassware should be used.
- gas cylinders should be replaced at 700-1400 kPa.
- laboratory personnel should have appropriate training in analytical laboratory procedures, and in the particular analysis for which they are responsible.

6.4.7.2 Calibration Control

Statistical control over calibration can be confirmed between runs by means of two control standards, A and B, and within-run by means of blanks and midrange standards.

a) **Between-run Calibration Control** Two control standards, A and B, can be used to analyse and control between-run changes in calibration, once at the beginning of each analytical run. These standards are made up and maintained independently of the calibration standards and are normally chosen to be about 80% and 20% of full-scale, respectively. Results are accumulated over many runs and the sums (A + B) and differences (A - B) are plotted on control charts. During a specific run, a significant change in the sum (A + B) from the historical mean implies that a significant change in intercept has occurred, other factors remaining constant. A significant change in the difference (A - B) implies a significant change of slope, other factors remaining constant. Control and warning limits for A - B are calculated for the mean and the standard deviation of the population of differences:

Upper and Lower Warning Limits (UWL, LWL) = $X_{A\text{-}B} \pm 2\ SD_{A\text{-}B}$

Upper and Lower Control Limits (UCL, LCL) = $X_{A\text{-}B} \pm 3\ SD_{A\text{-}B}$

Control and Warning limits for A + B are similarly calculated using the same standard deviation:

$$UWL / LWL = X_{A+B} \pm 2\ SD_{A-B}$$

$$UCL / LCL = X_{A+B} \pm 3\ SD_{A-B}$$

The run should not proceed until it is shown that A + B and A - B are within control limits. Control limits should not exceed ± 5% of the average value for A+B and A-B.

b) **Within-run Calibration Control (Inorganic Analyses)** Within-run changes in calibration attributable to slope and baseline drift should be checked at regular intervals. This can be accomplished by use of a mid-range standard and reagent blank run after every 20 samples. Control limits should be established by each laboratory for each procedure. The drift should not exceed 10%. If a greater drift is detected, the analysis should be stopped, the instrument re-calibrated, and samples run after the last acceptable check sample and blank are reanalysed.

6.4.7.3 Precision

Precision is the degree of variation among individual measurements of the same variable using a specific analytical method and is usually expressed as the standard deviation of replicates (U.S. EPA 1990). Statistical control of analytical precision is maintained by analysing within-run duplicates at a frequency of at least 10%. Laboratory duplicates are separate aliquots split in the laboratory from a single sample.

The absolute difference between within-run duplicates is compared to a control limit determined from historical data. To obtain these control limits, the results of duplicate analyses are accumulated over many runs and sorted according to concentration ranges. Convenient concentration ranges are 0 to 20%, 20 to 50%, and 50 to 100% of full scale (King 1976). Within each concentration range, control limits for the absolute difference between within-run duplicates is determined from the formula:

$$UCL = D_4 \times R$$

where D_4 is a statistical factor and R is the mean difference between duplicates, 3.267 (ASTM 1986; Taylor 1987).

If the difference between laboratory duplicate analyses exceeds the upper control limit, the situation should be evaluated to determine the most appropriate corrective action.

6.4.7.4 Accuracy and Bias

Accuracy is the degree of agreement between an observed value and the true value as determined by analysis of an accepted reference material (U.S. EPA 1990c). The converse of accuracy is the degree of systematic error in the analysis or bias. Accuracy is controlled by means of method blanks and certified reference materials.

a) **Method Blanks** A method blank is an aliquot of reagent water equivalent in volume to the samples being processed and run in exactly the same manner as the samples. The method blank quantifies the level of contamination introduced to the samples during sample processing and analysis. Method blanks should be analysed at a frequency of 10% or 1/run, charted, and controlled at ± 2 SD (warning limits) and ± 3 SD (control limits). If a method blank is judged out of control and contaminated, those samples processed with the blanks and greater than the detection limit should be repeated for the variable(s) affected. In general, a method blank is considered free of contamination if the analysis yields results less than the detection limit or less than 0.1 times the level found in all associated samples (CAEAL 1991).

b) **Standard Reference Materials** Standard reference materials (SRMs) are samples available in different matrices that have been extensively analysed by several laboratories and have concentrations certified by standard-setting organisations such as the National Institute of Science and Technology, the U.S. EPA, the National Water Research Institute of Environment Canada and the National Research Council. When available, an SRM should be analysed at a frequency of 5% or 1/run. The matrix and concentration of the SRM should be as close as possible to the samples being analysed. The results of SRMs should be accumulated and control and warning limits determined as ±3 SD and ±2 SD, respectively.

6.4.7.5 Recovery

Recovery of the analyte over the entire analytical process is determined from matrix spikes, spiked blanks, and surrogate spikes.

a) **Matrix Spike** A matrix spike is a separate aliquot of a randomly chosen sample to which is added all the analytes of interest before processing of the sample. Analysis of a matrix spike gives an indication of the recovery efficiency obtained for the matrix particular to that sample. The sample should be spiked with all the analytes of interest at a concentration as close as possible to that concentration giving a response equal to the mid-level calibration standard. The spiking solution should be prepared from a stock source separate from that used for calibration. The recommended distribution of matrix spikes is 10% or 1/run. One method to calculate recovery is:

$$\% \text{ Recovery} = \frac{\text{Measured Conc. - Unspiked Conc.}}{\text{Spike Amount}} \times 100$$

The results of matrix spikes should be plotted on separate control charts for each matrix. In-house limits should be set on the basis of ± 3 SD on a minimum of 10 data points. In multi-parameter analyses, at least 90% of the analytes should have recoveries within the specified limits. Recoveries for inorganic analytes should fall within 75 to 125%. Recoveries for organic variables should fall within the limits specified in Table 4 of CAEAL (1991). If a matrix spike does not meet these criteria, the spike should be repeated. If the recoveries do not meet the criteria in the repeat analysis and there are no indications of other problems with the analysis, a matrix effect should be noted and reported.

b) **Spiked Method Blank** The spiked method blank is a separate aliquot of the same reagent water used for the method blank that is spiked with the compound of interest at a concentration as close as possible to the concentration of the mid-level calibration standard. The spiked method blank gives an indication of the reliability of a method without the matrix effects of real samples. The spiked method blank should be processed with and in the same manner as the samples. As with the matrix spike, the spiking solution should be prepared from stocks separate from those used for calibration. One method to calculate recovery is

$$\% \text{ Recovery} = \frac{\text{SMB - MB}}{\text{Spike Amount}} \times 100$$

where: SMB = Concentration measured in spiked method blank
MB = Concentration measured in the method blank

In-house recovery limits should be calculated for the spiked method blank based on ± 3 SD and a minimum of 10 data points. Recoveries for inorganic analyses should fall between 75% and 125%. Recoveries for organic variables should fall within 70% and 120%. If a spiked blank recovery does not meet the criteria established, the spike should be repeated. If the spike still does not recover, the samples related to the spike should be repeated. If insufficient sample remains for a repeat analysis, the results should be reported and flagged as suspect with an explanation.

6.4.7.6 Detection Limits

Detection limits should be reported as the method detection limit (MDL) as described by U.S. EPA (1984). The MDL is defined as the minimum quantity of an analyte that should be observed to justify the claim to have detected the analyte with a specified risk (normally 5% or 1%) of making a false detection. One method to calculate the MDL is from the standard deviation of the analysis at the lowest concentration range:

$$MDL = t_{0.05\ n-1} \times S \qquad (1)$$

where: $t_{0.05,\ n-1}$ is the one tailed value of Student's t for a 5% risk of false detection, n-1 degrees of freedom, and S is the standard deviation.

Ideally the standard deviation is calculated from low level replicate analysis on real samples having the same or similar sample matrix as the samples under consideration. This standard deviation can be calculated from a minimum of seven replicates in the same run using the standard statistical formula (U.S. EPA 1984). However, it is preferable to calculate S from between-within-run replicate pairs accumulated over many runs.

The standard deviation of low-level replicate pairs accumulated over a large number of analytical runs is:

$$S = \frac{\Sigma S\ D^2}{2n} \qquad (2)$$

where D is the individual replicate difference and n is the number of replicate pairs. A minimum of 40 replicate pairs is recommended (OMOE 1988). The value of either standard deviation is then entered in equation (1) to calculate the MDL.

6.4.7.7 Data Reporting Conventions

Established protocols for rounding-off analytical results should be followed. If too many figures are rounded-off before reporting, information is lost and real differences in the concentrations of samples from different locations or occasions may be concealed. Quality control procedures may be on a coarser basis than is desirable, or necessary, with the result that values of the mean, standard deviation, or other statistics of a set of results may be biased. Conversely, when too many significant figures are reported, relatively small, statistically insignificant differences may appear falsely large (Hunt and Wilson 1986).

The standard deviation of the analysis is the preferred criterion for deciding the number of significant figures (King 1989). The process of rounding-off should ensure retention of the digit that is in the same decimal position as the most significant digit in the calculated standard deviation. For example, if the analysis provides a value such as 12.345 and the calculated standard deviation based on within-run replicate analysis at this concentration level is 0.32, the result should be truncated to 12.3.

6.4.7.8 Analytical Precision and Accuracy

Precision is the degree of agreement among replicate analysis of a sample, usually expressed as the standard deviation. Reproducibility is the closeness of agreement between the results of measurement of the same parameter carried out under changed conditions of measurement. Reproducibility is the standard deviation obtained measuring the same sample in different analytical runs and is called between run precision. Between run precision includes variability due to calibration on different days, instrument drift and many other factors.

Precision is affected by random errors and is a measurable and controllable parameter. Precision should be estimated for all analyses by processing separate sample aliquots through the entire analytical method. A laboratory should monitor their precision and be able to report precision using several days of data. For metals and most other parameters, precision should be within 10%. For total suspended solids the precision should be within 15% at concentrations greater that 10 times the MDL. For pH, precision should be within ± 0.1 pH unit.

Accuracy is the combination of bias and precision of an analytical method which reflects the closeness of a measured value to the true value of a sample. Bias is a systematic error caused by something in the measuring system resulting in the data being high or low. Bias can be caused by a number of factors including contamination, mechanical losses, blanks, spectral interference, calibration errors or the influence of different operators. Accuracy is measured as per cent recovery of known concentrations such as certified reference materials, spiked samples or reference samples prepared by the laboratory and analysed as samples.

Whether data is considered accurate or inaccurate is relative to the final use of the data. A laboratory should monitor their accuracy and be able to report it using several days of data. For metals and most other parameters accuracy should be within 10%. For total suspended solids, accuracy should be within 15% at concentrations greater that 10 times the MDL. For pH, accuracy should be within ± 0.1 pH unit (MMER Schedule 3).

6.4.7.9 Instrument Maintenance

ICP/MS Tuning and Performance: For analyses using mass spectrometry detection, instruments should be tuned to provide a mass calibration of the instrument (CAEAL 1991; U.S. EPA 1988). Instrument tuning is performed upon method set-up or as a corrective action. CAEAL (1991) recommends use of PFTBA (perfluorotributylamine, FC43) for tuning and has set specific performance criteria. Perfluorokerosene (PFK) is normally used to calibrate high resolution mass spectrometers. In addition, ICP/MS system performance checks using specific compounds should meet the performance requirements presented in CAEAL (1991).

For instrumental quality control the instrument check standard should be analysed once per 10 samples to determine if significant instrument drift has occurred. If agreement is not within ± 5% of the expected values (or within the established control limits) the analysis of samples should be discontinued, the problem corrected and the instrument recalibrated.

Instrument Response Checks: Instrument response (absorption, counts, peak height, etc.) can be checked before calibration or analysis by comparing the response of a check sample used on an ongoing basis, to historical data. Instrument response can be charted in control charts with a warning limit set at ± 2 standard deviation and an out-of-control limit of ± 3 standard deviations.

Instrument Calibration Requirements: Calibration is recommended for all inorganic methods with a minimum of three concentration levels (5 recommended) of standards, plus a blank, so that all samples fall within the calibrated range (CAEAL 1991). In cases where instrument software limits the calibration to two points, the calibration range should not exceed the linear range of the instrument. Unless it is demonstrated that no bias is introduced, standards should be matched to the matrix of the samples.

6.4.8 Quality Assurance in the Laboratory

Quality assurance (QA) encompasses a wide range of internal and external management and technical practices designed to ensure data of known quality commensurate with the intended use of the data.

External quality assurance activities include participation in relevant inter-laboratory comparisons and audits by outside agencies. Outside audits may be based on performance in analysis of standard reference materials, or on general review of practices as indicated by documentation of sampling, analytical and QA/QC procedures, test results, and supporting data.

6.4.9 Recording and Reporting of QA/QC Information

6.4.9.1 Documentation

Documentation of all aspects of the analysis is recommended to confirm the quality and reliability of the analytical results. Storage of sample results and data associated with the analyses in hard copy or on computer file (with back-up) is recommended for at least six years after analysis. Information for each sample or batch of samples information is recommended on the following:

a) **Method Detection Limits (MDLs)** If published MDLs are different from the laboratory-determined MDLs (due to interference, dilutions, etc.), this should be recorded.

b) **Sample Storage Times** Records should be kept on the sampling date, date of receipt, date of sample preparation, and date of analysis. This information is normally handled as part of the sample tracking process.

c) **Instrument Performance and Maintenance** A log should be kept of instrument performance including records of tuning and instrument response. Maintenance or service records should be kept for each instrument.

d) **Quality Control Samples** Records of duplicate analyses, blanks, spiked blank recoveries, surrogate recoveries, matrix spike recoveries, results from certified reference materials, and records of calibration and calibration checks should be maintained.

e) **Sample Reception, Preparation, and Analysis** All anomalies in delivery, storage, condition, preparation, and analysis of samples should be recorded. These include any deviations from standard operating procedures.

6.4.9.2 Reporting of QA/QC Information

Analytical results are reported as a test or analysis report and should include all relevant data needed to assess the validity of the data; including QA/QC components. Items that should appear in a data report include:

- the report should be accurate, clear, unambiguous and objective
- report needs a title (Test Report, Report of Analysis, Quality Report)
- name, address and location of laboratory and location where tested
- unique identification of the report so it can be traced easily (serial number, group number)

- name and address of client
- identification or description of the sample tested
- condition of the test item (unpreserved, leaking bottle, where relevant)
- date of sample receipt; date of report
- identification of the analysis method and description of any non-standard tests
- reference to sample date and sampling method (grab sample, time proportioned composite sample etc.)
- deviations from the usual test method (filtering, pH adjustment, standard addition etc.)
- the analytical result with units clearly identified
- statement indicating whether or not the results were corrected for blanks.
- quality control data
- identify if result is qualified (did not pass QC tests, sample size too small etc.)
- signature of accountable person and date authorised
- subcontractors clearly identified
- updates or corrections to reports clearly identified
- should notify clients if new information invalidates reports already issued

Data below the analytical detection limit should be clearly reported as such along with the applicable MDL for that sample.

6.5 Effluent Characterization and Water Quality Monitoring during Investigation of Cause

Investigation of Cause occurs when a mine related effect has been identified in the fish population, fish tissue or benthic invertebrate community and previous monitoring has failed to provide a satisfactory explanation of the cause of the effect. At sites where there are effects on fish or the benthic invertebrate community, there are a number of effluent characterization and water quality monitoring techniques available which may help in understanding the nature and cause of effects. In addition, effluent characterization of additional mine related contaminants from other sources, particularly non-point sources (those that were not required under MMER), may be appropriate at some sites as part of Investigation of Cause. In some cases, these techniques could also provide valuable information during Focussed Monitoring, as supporting environmental data, when determining the extent and magnitude of effects observed in fish or the benthic invertebrate community. The techniques described below are recommended, and mines may use individual techniques or combinations of techniques, as appropriate, to address site-specific questions.

6.5.1 Measurement of Dissolved Metals during Investigation of Cause

When an effect on fish or benthic invertebrates is observed during Initial, Periodic or Focussed Monitoring, and total metal measurements in the exposure area are elevated, then measurement of dissolved metals may be warranted as part of Investigation of Cause. The measurement of both total and dissolved metals on effluent and exposure area samples in Investigation of Cause would assist in determining which metals might be causing or contributing to the observed effects.

The rationale for measuring dissolved metals during Investigation of Cause is based on the theory that it is the metals in the dissolved fraction, particularly the free metal ions, that are the most bioavailable to aquatic organisms. There is qualitative evidence to support this theory, especially in defined synthetic media (i.e. laboratory bioassays). However, this relationship appears to break down in natural waters, particularly in the presence of natural dissolved organic matter (ESG 1999).

The normal procedure for measuring dissolved metal concentrations involves the immediate filtration of the raw water sample through a 0.45 micron filter and then preserving the filtered sample with nitric acid to a pH of less than 2.0 to keep the dissolved metals in solution until the analysis. There are differences in the filter size used but the 0.45 micron filter is the most common. There are concerns that certain amounts of colloid bound metals, which are not really dissolved, can pass through this size of a filter membrane and some researchers recommend filter pore sizes as small as 0.1 microns to minimise this concern (EVS 1997).

In a report by Hall (1998) three filter systems were evaluated; syringe, in-line, and vacuum filters. Most of the filters tested came from two leading manufacturers; Gelman and Millipore. The recommendations from this study were as follows:

- overall, optimum performance was achieved with the ion chromatography Acrodisc syringe filter from Gelman and the Sterivex syringe filter capsule from Millipore. Nylon membranes should be avoided as they are slow and do not show superior contaminant characteristics. For samples high in particulate matter, the Millex LS 5μ m syringe prefilter is recommended.

- the Millicup bottle topped with a Durapore membrane (vacuum system); the in-line Gelman AquaPrep with Thermopor membrane; and the AquaPrep 250with Supor membrane are acceptable. With sufficient rinsing, these systems could also be employed: the Gelman syringe GHP Acrodisc; the Millipore Millex HV syringe; and the in-line Gelman groundwater capsule.

- for the filtration of acidic samples, the following systems should be avoided: the Gelman syringe nylon Acrodisc; the Millipore all-glass vacuum; the Gelman groundwater capsule; and the Gelman AquaPrep 250.

The above recommendations are based upon the contamination contributed by these filters and their ease of use. However, before a system is chosen it is important to evaluate its propensity to become clogged with sample water.

Anyone planning on completing dissolved metals analyses is strongly encouraged to consult Hall (1998) prior to sample collection.

6.5.2 Measurement of Reagents and Reagent by-Products used in Processing

A wide range of chemical reagents are used in the ore beneficiation processes, and these reagents, or the by-products of the decomposition of these reagents, may occur in mine effluent (EAG, 1992). As a result, reagents and reagent by-products may occur in receiving environments. Analyses for reagents and reagent by-products may be helpful in cases where observed effects cannot be attributed to metals and other parameters monitored regularly as part of the EEM program at a site.

6.5.3 Metal speciation for metals of concern

The form in which a metal (or metalloid) occurs (i.e. the speciation of the metal) can have a significant affect on the bioavailability and toxicity of that metal to aquatic organisms (Tessier and Turner 1995; Stumm and Morgan 1996). As a result, understanding the speciation of contaminants of concern can be important to understanding the nature and causes of effects on aquatic ecosystems.

Metal speciation can be measured through direct analysis, or estimated with modelling tools. Two of the most important factors affecting metal speciation are pH and oxidation state. Thus, no matter which method is to be used to assess metal speciation, it is important to take accurate field measurements of pH and dissolved oxygen. In addition, if samples are to be analysed for metal speciation, it is important to ensure that proper sample collection, preservation and storage procedures are followed.

6.5.3.1 Chemical Analysis of Metal Speciation

The speciation, or at least valency, of some metals in water samples can be determined analytically. It is not known how widespread commercial availability of this type of analysis is. Analytical methods include methods using ion-exchange, and electrochemistry. Measurement techniques for the speciation of metals in aqueous solutions have been comprehensively reviewed by Tessier and Turner (1995).

If samples are to be submitted for analysis for speciation then extra care should be taken in sample collection, handling and storage, since changes in factors such as the dissolved oxygen content of the sample could result in speciation of the sample changing significantly between the time of collection and the time of analysis. It is strongly recommended that the laboratory that will be completing the analyses be contacted in advance of sample collection to identify appropriate procedures.

6.5.3.2 Chemical Speciation Models

See also EVS Environmental Consultants (1997).

A number of computer programs are available for the modelling of chemical speciation of metals in water, for example:

MINTEQA2	(Allison et al. 1991)
MINEQL$^+$	(Schecher and McAvoy 1992)
WHAM	(Kramer et al. 1996, Hare and Tessier 1996)

Since metal speciation is dependant on a number of factors, all of these programs require data on: pH, alkalinity, hardness of calcium concentration, ionic strength, total dissolved solids, dissolved organic carbon, total dissolved metal, unique anthropogenic inputs such as EDTA, and concentrations of aluminium and iron (III) (EVS 1997).

Potential sources of error include (EVS 1997):

- extrapolations to situations which have not been sufficiently parameterized, such as cases where the pH is above 7.
- poor data, which can result in more variation for some parameters than for others
- parameter uncertainty, such as variations in thermodynamic data in program databases.
- difficulties in modelling metal interactions with natural particulate matter and dissolved organic matter.

These programs can be used to predict the forms and concentrations of metals in effluent or water, and can be used to predict toxicity (EVS 1997).

6.5.4 Measurement of Chlorophyll A and/or Phaeo-pigment

The measurement of chlorophyll a, chlorophyll b, and phytoplankton biomass may be helpful in cases where upstream discharges of nutrients represent confounding influences. The concentration of photosynthetic pigments is often used to estimate phytoplankton biomass.

Chlorophyll a data are used as indicators of phytoplankton biomass, and therefore as a qualitative indicator of the productivity of an aquatic ecosystem. Chlorophyll a can be measured using spectrophotometry, fluorometry, or liquid chromatography. It is reported in mg/m^3, either as chlorophyll a, or converted to algal biomass using a conversion factor of 67 for 1.5% chlorophyll in algal tissues.

When measuring chlorophyll a, it is necessary to take into account chlorophyll b (and sometimes c); and corrections are routinely made to take this into account. Depending on the ecological setting (the supply of decomposing vegetation), varying amounts of chlorophyll degradation products are also present (chlorophyllides, pheophorbides and pheophytins); corrections can be made to take these components into account, and it is also possible to report separately the chlorophyll degradation products. A correction can also be made for turbidity – particulate matter in suspension. More information can be found on the analytical methods in Standard Methods (method 10200 H).

Phytoplankton biomass tends to be patchy in its distribution, and varies with time as blooms develop and waters stratify. It is more suitable for lake than for river settings, although evidence for significance phytoplankton biomass in riverine samples is nevertheless an indicator that nutrients reflecting trophic status upstream may be, and probably are, relevant in assessing nutrient (as opposed to contaminant) influences on the downstream structure of benthic communities, zooplankton and presumably fish as well.

6.5.5 Monitoring of Flow and Loadings in the Exposure Area

During Investigation of Cause measuring water flow in the exposure area could help with data interpretation and understanding of the dilution ratio, mixing, mass balance, fate and effects of contaminants and causal relationships. It may be useful to evaluate the degree of exposure of communities to the mining effluent and its contaminants over a longer period of time and in a variety of conditions.

Assessments of loading rates are essential to mass balance calculations. Assuming complete mixing, the principal statement for mass balance at an industrial outfall is:

Mass or loading rate of substance upstream + Mass rate added by outfall
= Mass rate of substance immediately downstream of the outfall.

Since the loading rate is the product of flow and concentration, the mass balance is given by:

$$Q_0 s_0 + Q_e s_e \quad = \quad Qs$$

Where Q_0, Q_e, and Q= flow rate upstream, of effluent, and downstream, respectively; and

s_0, s_e, and s = concentration upstream, in effluent, and downstream, respectively.

A similar statement can be made for the balance of flows:

$$Q_0 + Q_e = Q$$

Upstream conditions of flow and concentrations are often known or can be readily measured. Effluent characteristics are usually known in greater detail. Solving for the downstream concentration, s, yields:

$$s = \frac{Q_o s_o + Q_e s_e}{Q}$$

The downstream concentration is therefore dependent on the upstream and downstream flows and the concentration of upstream and effluent chemical inputs. If the upstream concentration is zero, then the downstream concentration, s, will equal the effluent concentration reduced by the ratio of effluent flow to total river flow. This is commonly known as a dilution effect.

The above equations apply to the simplest scenario immediately downstream of the outfall. Mass balance computations become more complex as one moves further downstream of the outfall and contributing factors of tributaries, multiple point sources, distributed sources, as well as dispersion, chemical settling, and chemical decay effects are considered. Useful references for formulating appropriate mass balance equations include Thomann and Mueller (1987); McCutcheon and French (1989) and Henderson-Sellers and French (1991).

6.6 References

Allison JD, Brown DS and Novo-Gradac KJ. 1991. MINTEQA2/PRODEFA2, a geochemical assessment model for environmental systems. Version 3.0 User's Manual. U.S. Environmental Protection Agency, Washington, D.C.

APHA. 1995. Standard Methods for Examination of Water and Wastewater. 19th edition, published by the American Public Health Association (APHA).

APHA. 1999. Standard Methods for Examination of Water and Wastewater. 20th edition, published by the American Public Health Association (APHA).

ASTM. 1986. Manual on Presentation of Data and Control Chart Analysis. Committee E-11 on Statistical Methods. ASTM Special Technical Publication 15D. American Society for Testing and Materials, Philadelphia.

CAEAL (Canadian Association of Environmental Analytical Laboratories). 1991. Code of Practice and QA Manual for Laboratory Analysis of Sewage Treatment Effluent in Support of the MISA Program; Draft report prepared for CAEAL and the Ontario Ministry of the Environment by Zenon Environmental Laboratories.

Ecological Services Group (ESG). 1998. Technical Evaluation of Toxicity Identification Evaluations (TIEs) and Toxicity Reduction Evaluations. Report prepared for AETE program, Natural Resources Canada. CANMET, Ottawa, Ontario.

Ecological Services Group (ESG). 1999. AETE Synthesis Report of Selected Technologies for Cost-Effective Environmental Monitoring of Mine Effluent Impacts in Canada (AETE Project No. 4.1.4). Report for AETE program, CANMET, Natural Resources Canada, Ottawa, Ontario

ELAP (Environmental Laboratory Approval Program). 1988. Environmental Laboratory Approval Program Certification Manual. New York State Department of Health.

Environment Canada. 2001. Guidance Document for the Sampling and Analysis of Mine Effluents. Report EPS2/MM/5. Published by Environment Canada.

Environment Canada. 1999. Metal Mining Environmental Effects Monitoring Requirements. EEM/1999/1.

Environment Canada. 1998. Pulp and paper EEM technical guidance document. EEM/1998/1, April, 1998.

EVS Environmental Consultants. 1997. Technical evaluation: water quality and biological effects. Report for AETE program, CANMET, Natural Resources Canada, Ottawa, Ontario.

Graham J. 1997. Standard Method for the Determination of Total Thiosalt. Mining and Mineral Sciences Laboratories, Divisional Test Report 1997-067, Natural Resources Canada, Ottawa, ON.

Hall GEM. 1998. Cost-effective protocols for the collection, filtration and preservation of surface waters for detection of metals and metalloids at ppb and ppt levels. Report for AETE program, CANMET, Natural Resources Canada, Ottawa, Ontario.

Hare L and Tessier A. 1996. Predicting animal cadmium concentrations in lakes. Nature 380:430-432.

Henderson-Sellers B., and French RH.. 1991. Water Quality Modelling Volume IV: Decision Support Techniques for Lakes and Reservoirs. CRC Press, Boca Raton, Florida.

Hunt DTE and Wilson AL. 1986. The Chemical Analysis of Water; General Principles and Techniques, 2nd Ed. The Royal Society of Chemistry, London.

Kramer JR, Davidson W, Godtfredsen KL, Meyer JS, Perdue M, Tipping E, Van de Meent D and Westall J. 1996. Chemical speciation and metal toxicity in surface waters. Pp5.1-5.17. In: Reassessment of metals criteria for aquatic life protection. Bergman H, Dorward-King E, Allen H, DiToro D, Erickson R, Matice J and Reiley M (eds.). SETAC. Pensacola, FL.

McCutcheon SC and French RH. 1989. Water Quality Modelling Volume I: Transport and Surface Exchange in Rivers. CRC Press, Boca Raton, Florida.

McQuaker NR. 1999. Technical Evaluation on Water Quality Design and Analysis (AETE Project No. 3.1.1) Draft. Report for AETE program, CANMET, Natural Resources Canada, Ottawa, Ontario.

Nielson LA and Johnson DL. 1983. Fisheries Techniques. The American Fisheries Society, Southern Printing Company, Virginia. 468 p.

NLET (National Laboratory for Environmental Testing). 1994a. Manual of Analytical Methods, Vol. 1. Major Ions and Nutrients. NLET, NWRI, Environment Canada, Burlington, Ontario.

NLET (National Laboratory for Environmental Testing). 1994b. Manual of Analytical Methods, Vol. 2. Trace Metals. NLET, NWRI, Environment Canada, Burlington, Ontario.

Ontario Ministry of the Environment (OMOE). 1988. Estimation of Analytical Detection Limits (MDL) Report by the Ontario Ministry of the Environment ISBN-0-7729-4117-3. Schecher WD and McAvoy DC. 1992. MINEQL+: A software environment for chemical equilibrium modeling. Computers, Environment and Urban Systems 16:65-76.

Stumm W and Morgan JJ. 1996. Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters, 3rd. ed. Wiley.

Taylor JK. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers Inc., Chelsea, Michigan.

Tessier A and Turner DR (eds.). 1995. Metal Speciation and Bio-availability in Aquatic Systems. Wiley.

Thomann RT and Mueller JA. 1987. Principles of Surface Water Quality Modelling and Control. Harper Collins Publishers Inc., New York, New York.

U.S. Environmental Protection Agency (EPA). 1984. Definition and Procedure for the Determination of the Method Detection Limit – Revision 1.11; Appendix B to Part 136. Federal Register Vol. 49, No. 209, Oct. 26, 1984, Part VI, 40 CFR Part 136.

U.S. Environmental Protection Agency (EPA). 1990c. Proposed Glossary of Quality Assurance Related Terms. QAMS RD-680.

U.S. Environmental Protection Agency (EPA). 1991. EPA 200 Series – Methods for the Determination of Metals in Environmental Samples, EPA/600/4-91-010, June, 1991.

U.S. Environmental Protection Agency (EPA). 1994. EPA 200 Series – Methods for the Determination of Metals in Environmental Samples. Supplement 1, EPA/600/R-94-111, May 1994.

Appendix 6.1: Justifications for Parameters for Effluent Characterization and Water Quality Monitoring

6.7 Deleterious Substances and pH *as per* Schedule 3 of the MMER

Arsenic:

- AQUAMIN recommended that arsenic be measured in effluent characterization;
- Arsenic can occur in effluent from a wide range of mine types, including gold, base metal, and uranium, and can be expected to occur across Canada;
- The MISA draft development document[1] stated that arsenic was found in 26% of the metal mine effluents sampled, with an average concentration of 0.036 mg/L;
- Arsenic can bioaccumulate in fish and is known to be toxic to aquatic organisms; and,
- The Canadian Water Quality Guideline[2] for arsenic for the protection of freshwater aquatic life is 0.005 mg/L.

Copper:

- AQUAMIN recommended that copper be measured in effluent characterization;
- Copper can occur in effluent from a wide range of mine types, particularly gold and base metal, and can be expected to occur across Canada;
- The MISA draft development document reported that for 39 Ontario effluent streams sampled for 12 months, the average copper concentration was 0.07 mg/L;
- Copper is known to be toxic to aquatic organisms; and,
- The Canadian Water Quality Guideline for copper for the protection of freshwater aquatic life ranges from 0.002 to 0.004 mg/L, depending on the water hardness,

Lead:

- AQUAMIN recommended that lead be measured in effluent characterization;
- Lead can occur in effluent from a wide range of mine types, particularly base metal mines, and can be expected to occur across Canada;
- The MISA draft development document stated that lead was found in 20% of the metal mining effluents sampled. The average lead concentration in the sampled effluent was 0.02 mg/L;
- Lead is known to be toxic to aquatic organisms; and,

---

[1] Note: throughout this document, the MISA draft development document is referred to. This document summarises data from one year of comprehensive effluent characterization at mines across Ontario in the early 1990s. This data is not representative of the current state of effluent quality from mines across Canada. However, since no database exists which is more recent, this information is reported to provide some indications of the potential occurrence of contaminants in mine effluent.

[2] Guidelines are quoted throughout this document to provide a preliminary assessment of effluent concentrations, with respect to ecological significance, and potential for environmental effects in the receiving environment.

- The Canadian Water Quality Guideline for lead for the protection of freshwater aquatic life ranges from 0.001 to 0.007 mg/L, depending on the water hardness.

Nickel:

- AQUAMIN recommended that nickel be measured in effluent characterization;
- Nickel can occur in effluent from a wide range of mine types, particularly base metal and uranium mines, and can be expected to occur across Canada;
- The MISA draft development document stated that nickel was found in 68% of the metal mine effluents sampled, with an average concentration of 0.14 mg/L;
- Nickel is known to be toxic to aquatic organisms; and,
- The Canadian Water Quality Guideline for nickel for the protection of freshwater aquatic life ranges from 0.025 to 0.150 mg/L, depending on the water hardness.

pH:

- AQUAMIN recommended that pH be measured in effluent characterization;
- pH extremes can occur in effluent from a wide range of mine types and can be expected to occur across Canada;
- pH often determines the solubility of metal species and, therefore, is linked to the toxicity of the effluent;
- Extremes of pH are known to be toxic to aquatic organisms; and,
- The Canadian Water Quality Guideline for pH for the protection of freshwater aquatic life is 6.5 to 9.0.

Radium 226:

-
- AQUAMIN recommended that radium 226 be measured in effluent characterization;
- Radium 226 occurs primarily in effluent from uranium mines, and does not occur across Canada; and,
- There is no Canadian Water Quality Guideline for radium 226.

Total Cyanide:

- Cyanide is used as a process reagent at most gold mines and some base metal mines;

- AQUAMIN recommended that cyanide be measured in effluent characterization, for mines that use cyanide as a process reagent;
- The MISA draft development document stated that cyanide was found in 54% of the metal mine effluents sampled, with an average concentration of 0.084 mg/L in gold mining effluent and 0.006 mg/L in base metal mining effluent;
- Cyanide is known to be toxic to aquatic organisms; and,
- The Canadian Water Quality Guideline for free cyanide for the protection of freshwater aquatic life is 0.005 mg/L.

Total Suspended Solids:

- AQUAMIN recommended that total suspended solids be measured in effluent characterization;
- Suspended solids can occur in effluents from all mine types, and occur across Canada;
- The MISA draft development document stated that suspended matter was found in 80% of the metal mine effluents sampled, with an average concentration of 7 mg/L; and,
- Suspended solids can kill fish by clogging their gills, and can effect fish habitat by smothering fish habitat, contaminating sediments, or reducing light penetration in receiving waters.

Zinc:

- AQUAMIN recommended that zinc be measured in effluent characterization;
- Zinc can occur in effluent from a wide range of mine types, particularly base metal mines, and can be expected to occur across Canada;
- The MISA draft development document stated that zinc was found in 76% of the metal mine effluents sampled, with an average concentration of 0.07 mg/L;
- Zinc is known to be toxic to aquatic organisms; and,
- The Canadian Water Quality Guideline for zinc for the protection of freshwater aquatic life is 0.030 mg/L.

6.8 Parameters Required in Schedule 5 of MMER for Effluent Characterization and Water Quality Monitoring

Alkalinity:

- Alkalinity is a measure of the buffering capacity of water, and gives an indication of how sensitive water is to changes in pH; and,
- Alkalinity is a factor affecting the fate and bioavailability of metals.

Aluminium:

- AQUAMIN recommended that aluminium be measured in effluent characterization;
- Aluminium occurs in a number of important rock forming minerals, and tailings pond effluents from a range of mine types may contain dissolved aluminium ions as well as chemically bound aluminium in the form of clays and other alumino-silicate mineral particles;
- The MISA draft development document stated that aluminium was present in 70% of the Ontario metal mining effluents sampled, with an average concentration of 0.20 mg/L;

- The draft development document also stated that dissolved aluminium is not a significant component of most metal mining effluents;
- The Canadian Water Quality Guideline for aluminium for the protection of freshwater aquatic life ranges from 0.005 to 0.100 mg/L, depending on the water hardness; and,
- Aluminium is toxic to aquatic organisms and its toxicity varies with pH.

Cadmium:

- AQUAMIN recommended that cadmium be measured in effluent characterization;
- Cadmium occurs in a relatively small range of ore types, but can be expected to occur at mines across Canada;
- The MISA draft development document states that cadmium was found in 12% of the metal mine effluents sampled, with an average concentration of 0.003 mg/L;
- Cadmium is known to be toxic to aquatic organisms and is bioaccumulative; and,
- The Canadian Water Quality Guideline for cadmium for the protection of freshwater aquatic life is 0.000017 mg/L.

Iron:

- AQUAMIN recommended that iron be measured in effluent characterization;
- Iron occurs in virtually all ore types, and occurs in mines across Canada;
- The MISA draft development document stated that iron was found in 100% of the metal mine effluents sampled, with an average concentration of 0.45 mg/L;
- Iron is toxic to aquatic organisms, and iron hydroxide precipitates can affect fish habitat;
- The Canadian Water Quality Guideline for iron for the protection of freshwater aquatic life is 0.30 mg/L; and,
- Iron can also have an important influence on the fate of other contaminants.

Nitrogen Compounds (ammonia and nitrate)

- AQUAMIN recommended that total ammonia be measured in effluent characterization;
- Nitrogen compounds are used in explosives in most mines, and residue from these explosives can result in nitrogen compounds occurring in effluent. Nitrogen compounds can also occur in effluent as a result of the breakdown of cyanide;
- During MISA development in Ontario, ammonia plus ammonium, Total Kjeldahl Nitrogen, and nitrate plus nitrite were measured in effluents;
- "Ammonia plus ammonium" was found in all 50 (100%) of the metal mining effluents sampled. The average concentration of "ammonia plus ammonium" measured in the metal mining sector was 1.4 mg/L in base metal mining effluents and 6.3 mg/L in gold mining effluents;
- "Total Kjeldahl nitrogen" was found in 96% of the metal mining effluents sampled. The average concentration of "total Kjeldahl nitrogen" measured in the metal mining sector was 8 mg/L;
- Nitrate plus nitrite" was found in 90% of the metal mining effluents sampled. The average concentration of "nitrate plus nitrite" measured in the metal mining sector was 8.8 mg/L;

- Nitrogen compounds can be toxic to aquatic organisms . In addition, nitrogen compounds are nutrients, and can lead to excessive plant growth. Excessive plant growth can lead to oxygen depletion, resulting in fish kills;
- The Canadian Water Quality Guideline for total ammonia for the protection of freshwater aquatic life ranges from 1.370 to 2.200 mg/L, depending in temperature and pH;
- The proposed interim Canadian Water Quality Guideline for nitrate is 13 mg/L in freshwater; and,
- The Canadian Water Quality Guideline for nitrite for the protection of freshwater aquatic life is 0.060 mg/L.

Mercury:

- AQUAMIN recommended that mercury be measured in effluent characterization;
- Mercury occurs in a range of rock types. It can occur at gold and silver mines, and less commonly at base metal mines, and is expected to occur across Canada;
- Mercury can come from a range of sources, including airborne transport, natural sources, and mine effluent. As a result, AQUAMIN concluded that it is often difficult to determine the source of mercury contamination within an aquatic environment. This was the basis for recommending the mercury be included in effluent characterization;
- The MISA draft development document stated that mercury was found in 14% of the metal mine effluents sampled, with an average concentration of 0.0002 mg/L;
- Mercury is toxic to aquatic organisms, and biomagnified within food chains; and,
- The Canadian Water Quality Guideline for mercury for the protection of freshwater aquatic life is 0.0001 mg/L.

Molybdenum:

- AQUAMIN recommended that molybdenum be measured in effluent characterization;
- Molybdenum occurs primarily in uranium ores, but may also occur in base metal ores and a small number of gold ores, and is not expected to occur across Canada;
- Molybdenum can be toxic to aquatic organisms, but its toxicity is not well understood;
- Molybdenum can also affect drinking water quality and cause molybdenosis in cattle; and,
- The Canadian Water Quality Guideline for molybdenum for the protection of freshwater aquatic life is 0.073 mg/L.

Hardness:

- Water hardness is a measure of cations, predominantly divalent cations, dissolved in water;
- Calcium and magnesium are the major contributors to hardness, and hardness can be calculated based on concentrations of these ions; and,
- Hardness is an important factor affecting the fate, bioavailability and toxicity of metals.

6.9 Required Parameters for Water Quality Monitoring Only

Dissolved oxygen:

- Dissolved oxygen can be measured in the field or in the lab; and,
- Dissolved oxygen is a factor affecting the fate and bioavailability of metals.

Temperature:
- Temperature changes can affect limnological properties of lakes, and can also affect aquatic organisms; and,
- the Canadian Water Quality Guideline for temperature for the protection of freshwater aquatic life states that thermal additions should not alter thermal stratification or turnover dates, exceed maximum weekly temperature averages, nor exceed maximum short-term temperatures.

6.10 Additional Site Specific Recommended Parameters for Effluent Characterization and Water Quality Monitoring

Calcium:
- AQUAMIN recommended that calcium be measured in effluent characterization;
- Calcium is an important cation in aquatic ecosystems, and may also occur in mine effluent, as a result of acid neutralisation using lime;
- Calcium discharges from mines may have effects on fish habitat;
- Calcium concentrations are essential to calculating hardness;
- Calcium is known to affect the toxicity of metals and/or other mine effluent parameters; and,
- Calcium data may assist in the interpretation of the potential impacts of metals and other parameters.

Chloride:
- The MISA draft development document stated that chloride was found in 80% of the metal mine effluents sampled, with an average concentration of 83 mg/L;
- Chloride can affect the toxicity of metals and/or other mine effluent parameters; and,
- Chloride data may assist in the interpretation of the potential impacts of metals and other parameters.

Conductivity:
- Conductivity is a measure of water's ability to conduct an electrical current;
- Conductivity can be measured in the field or in the lab; and,

- Conductivity gives an approximate measure of total dissolved solids, and can be used to identify the location of an effluent plume in freshwater environments.

Fluoride:

- AQUAMIN recommended that fluoride be measured in effluent characterization for operations where it is likely to be present;
- Fluoride occurs in a limited number of ore types, and is not expected to occur across Canada;
- Fluoride has been shown to bioaccumulate in fish bones, but its effects on aquatic organisms are not well understood;
- Fluoride is lethal to fish at concentrations ranging from 10 to 200 mg/L; and,
- Fluoride was rated "toxic" following Priority Substance List 1 assessment for "having an immediate or long-term harmful effect on the environment".

Magnesium:

- Magnesium is an important cation in aquatic ecosystems, and magnesium concentrations are essential to calculating hardness;
- Magnesium is known to affect the toxicity of metals and/or other mine effluent parameters; and,
- Magnesium data may assist in the interpretation of the potential impacts of metals and other parameters.

Manganese:

- AQUAMIN did not consider manganese for inclusion in effluent characterization;
- Manganese occurs in many ore types, and is expected to occur at mines across Canada Manganese makes up 0.1% of the earth's crust;
- There is no Canadian Water Quality Guideline for manganese for the protection of freshwater aquatic life;
- Manganese can have an important influence on the fate of other contaminants, specifically on the availability of other metals; and,
- Manganese is toxic to aquatic life but the factors that affect its toxicity are not well understood.

Potassium:

- Potassium is known to affect the toxicity of metals and/or other mine effluent parameters; and,
- Potassium data may assist in the interpretation of the potential impacts of metals and other parameters.

Selenium:

- AQUAMIN did not consider selenium for inclusion in effluent characterization;
- The MISA draft development document stated that selenium was found in 10% of the metal mine effluents sampled, with an average concentration of 0.007 mg/L;
- Selenium occurs most commonly in association with sulphur, and is not common in mine effluents;

- Selenium is toxic to aquatic organisms; and,
- The Canadian Water Quality Guideline for selenium for the protection of freshwater aquatic life is 0.001 mg/L.

Sodium:

- Sodium is known to affect the toxicity of metals and/or other mine effluent parameters; and,
- Sodium data may assist in the interpretation of the potential impacts of metals and other parameters.

Sulphate:

- Sulphate is an important anion in water;
- Mine effluents from mines with sulphide-bearing ore can be important sources of sulphate; and,
- The MISA draft development document stated that sulphate was found in 86% of the metal mine effluents sampled, with an average concentration of 644 mg/L.

Total phosphorus:

- AQUAMIN did not consider total phosphorus for inclusion in effluent characterization;
- The MISA draft development document stated that total phosphorus was found in 24% of the metal mine effluents sampled, with an average concentration of 0.1 mg/L;
- Total phosphorus may occur in mine effluent in particulate or dissolved form; and,
- Total phosphorus is a nutrient, and may lead to excessive plant growth.

Uranium:

- AQUAMIN did not consider uranium for inclusion in effluent characterization;
- Small amounts of uranium occur in many rock types in Canada, but uranium is expected to occur primarily in effluents from uranium mines;
- The MISA draft development document stated that uranium was found in 22% of the metal mine effluents sampled, with an average concentration of 0.06 mg/L; and,
- Uranium has been shown to bioaccumulate in fish.

6.10.1 Parameters Recommended for Effluent Characterization Only

Total thiosalts:

- AQUAMIN recommended that thiosalts be measured in effluent characterization (listed in Table 4.1, but not explicit in text) and also recommended that thiosalt impacts be monitored as part of EEM;
- Thiosalts are a group of metastable oxysulphur anions formed by partial oxidation of sulphide minerals during processing;
- Thiosalts have low toxicity, but in the aquatic environment the oxidation of thiosalts can lead to significant reductions in pH;
- Reductions in fish and benthic communities have been associated with thiosalts; and,

- Thiosalts have the potential to occur at any mine using flotation to separate sulphide minerals, but only a few mines in Canada have known problems associated with thiosalts.

6.10.2 Site specific Parameters Recommended for Water Quality Monitoring Only

Dissolved and total organic carbon:

- Dissolved and total organic carbon are important factors affecting the fate and bioavailability of metals.

Salinity:

- Salinity is an important parameter in marine environments, and may be a contaminant at some uranium mines.

Optical depth or Transparency:

- Optical depth is a field measurement of transmission of light through water as affected by water colour (dissolved constituents) and turbidity (particulate constituents);
- Optical depth is measured in the field using a Secchi disk or a turbidimeter; and,
- A low level of light transmission can reduce primary productivity in water and reduce the ability of predators to find prey.

Water Depth:

- Water depth can affect temperature, dissolved oxygen, and the degree of effluent dilution, all of which are modifying factors in effluent toxicity.

Table of contents

7. Sublethal Toxicity Testing 7-1

7.1 Introduction 7-1
- 7.1.1 Overview 7-1
- 7.1.2 Uses of Sublethal Toxicity Testing 7-1
- 7.1.3 Frequency of Sublethal Toxicity Testing 7-4
- 7.1.4 Sampling Locations 7-5
- 7.1.5 Reporting 7-5

7.2 Description of Freshwater and Marine Sublethal Toxicity Tests 7-6
- 7.2.1 Macrophyte Growth Inhibition Test Using *Lemna minor* 7-6
- 7.2.2 Algal Growth Inhibition Test Using *Selenastrum capricornutum* 7-7
- *7.2.3 Ceriodaphnia dubia* Reproduction and Survival Test 7-8
- 7.2.4 Fathead Minnow *(Pimephales promelas)* Growth and Survival Test 7-8
- 7.2.5 Rainbow Trout *(Oncorhynchus mykiss)* Embryo Development Test 7-9
- 7.2.6 Fertilization Assay Using Echinoids (Sea Urchins and Sand Dollars) 7-10
- 7.2.7 Larval Growth and Survival Assay Using Inland Silverside (*Menidia beryllina*) 7-11
- 7.2.8 Larval Growth and Survival Assay Using Topsmelt (*Atherinops affinis*). 7-12
- 7.2.9 Sexual Reproduction Assay Using the Red Macroalga (*Champia parvula)*7-13

7.3 Dilution Water Issues in Freshwater Sublethal Toxicity Testing 7-13
- 7.3.1 Dilution Water Selection 7-13
- 7.3.2 Collection, Shipment and Storage of Samples for Sublethal Toxicity Testing7-18
- 7.3.3 Test Organism Acclimation 7-20
- 7.3.4 Screening Tests 7-22
  - 7.3.4.1 Screening Test Impairment 7-22
- 7.3.5 Effluent or Effluent Exposed Surface Water Toxicity Tests 7-23

7.4 Quality Assurance and Quality Control (QA/QC) 7-24

7.5 Use of Sublethal Toxicity Testing in Resolving Confounding Influences 7-24
- **7.5.1 Microtox Chronic Test 7-26**

7.6 Investigation of Cause 7-27

7.7 References 7-28

List of tables

Table 7-1: Methodologies for effluent sublethal toxicity tests 7-3
Table 7-2: Specifications of the Environment Canada test methods and recommendations for collection, storage and use of site- collected dilution waters7-16
Table 7-3: Dilution/control water and corresponding effluent volumes for sublethal toxicity tests* 7-19
Table 7-4 Preparation of different hardness/alkalinities 7-21

7. Sublethal Toxicity Testing

7.1 Introduction

7.1.1 Overview

The purpose of sublethal toxicity testing is to answer the question: "Is there evidence from sublethal toxicity testing that mine effluent may affect fish, invertebrates and aquatic plants?" Acceptable sublethal toxicity data generated through the testing of an effluent from a specific discharge location over time can provide an indication of the degree of variability in effluent quality and temporal or seasonal trends.

To answer the above question mines will conduct a suite of four freshwater sublethal toxicity tests (fish, invertebrate, algae, and plant) or three marine sublethal toxicity tests (fish, invertebrate and algae), depending on the receiving environment type. The species chosen should be based on the relevance of the species to the local receiving environment. The sublethal toxicity methods to be used to conduct the tests are outlined in Table 7.1. Information on the relative sensitivity of the different sublethal toxicity tests can be found in ESG (1999). Effluent samples used for sublethal toxicity testing will be sub-samples of samples collected for Effluent Characterization (see section 6.1.1).

Following the first three years of the EEM program, adjustments to the sublethal toxicity testing requirement may be recommended after the completion of the national review of the entire EEM program. For the purpose of the national review, mines must complete a site-specific interpretation of the relationship between the sublethal toxicity testing results and the results of the fish survey and benthic invertebrate community survey, to be included in the interpretative report. The national review should be completed within a reasonable time, no later than two years following its commencement.

7.1.2 Uses of Sublethal Toxicity Testing

There are two potential uses of sublethal toxicity tests in the EEM program.

1) To estimate the potential for exposure area impacts and to estimate the potential geographic extent of the sublethal response using the results from toxicity tests conducted on mining effluent.
2) To measure changes in effluent quality over time as a result of changes in water treatment, site runoff management, acid rock drainage (ARD) mitigation measures or mine process changes.

It should be pointed out that the logic behind using sublethal toxicity data as part of the EEM program is to provide an estimate of the extent of potential effects on biological components of the exposure area, whether these components are being directly measured in the field or not.

Where sublethal toxicity data has an IC25 result (i.e., effluent concentration where a 25% inhibition is observed in the exposed test organisms) of less than 30%, it is recommended that mines calculate the geographic extent of the response in the exposure area and identify the zone where the concentration of effluent is comparable to the IC25 result. To complete this estimate data on effluent and receiving flow for the appropriate month would be needed. The estimation of the potential geographic extent can be effectively reported in map form.

Table 7-1: Methodologies for effluent sublethal toxicity tests

Test Description	Receiving Environment	Recommended Test Species	Methods
Fish early life stage Development tests	Marine	Inland silverside *(Menidia beryllina)* or Topsmelt (*Atherinops affinis*)	U.S.EPA 1994a and 1995
	Freshwater	Fathead minnow *(Pimephales promelas)*[1] or Rainbow trout *(Oncorhynchus mykiss)*	EC 1992a or EC 1998
Invertebrate Reproduction tests	Marine	Echinoids (sea urchins or sand dollars)	EC 1992b
	Freshwater	*Ceriodaphnia dubia*	EC 1992c
Plant and Algae Toxicity tests	Marine - algae	*Champia parvula*	U.S.EPA 1994a
	Freshwater - algae	*Selenastrum capricornutum*	EC 1992d or MEQ 1997[2]
	Freshwater - plant	*Lemna minor*	EC 1999a

1. Where fathead minnows are not an indigenous species, rainbow trout will be used according to EC 1998.
2. In some jurisdictions, the provincial testing requirements for *Selenastrum* may be the MEQ 1997. In order to harmonize with provincial requirements, this methodology would be acceptable for the EEM program.

NB: For all marine toxicity test procedures, it is recommended that the effluent salinity adjustment procedure by Environment Canada (2001) be followed. For all sublethal tests where the test organisms are purchased for sublethal toxicity testing, it is recommended the test organism importation of Environment Canada (1999b) be followed.

7.1.3 Frequency of Sublethal Toxicity Testing

Sublethal toxicity tests will be conducted twice each calendar year for the first three years. The first effluent sample will be collected, and sublethal toxicity testing conducted, no later than six months after the mine is subject to section 7 of the MMER *(Schedule 5, s. 6(1))*.

Following the first three years of the program, the mine may collect effluent, and conduct sublethal toxicity testing once per calendar year.

Sublethal toxicity testing may be conducted once a year for the first three years after a mine is subject to section 7 of the MMER, if the results of six sublethal toxicity tests conducted after December 31st, 1997 on a fish species, an invertebrate species and either an aquatic plant species or an algal species are submitted in a report to the authorization officer no later than six months after the mine becomes subject to section 7 of the MMER *(Schedule 5, s. 6(2))*.

In determining whether or not to use historical sublethal toxicity data as part of the EEM program, the mine should take into consideration:

- laboratory QA in the methods listed in Table 7.1;
- no less than three species tested (adequate to answer the question);
- age of the data (testing conducted after December 31st, 1997);
- the nature of mine operating conditions (e.g., are mine operations similar to the operations when the sublethal toxicity tests were conducted?); and,
- whether any of the sublethal toxicity testing was conducted on an effluent sample taken concurrent with fish and benthic invertebrate monitoring.

Collection of Samples

Effluent samples for sublethal toxicity testing will be taken from aliquots of effluent samples collected as part of Effluent Characterization (Chapter 6). *(Schedule 5, s. 5(2))*.

In choosing when effluent for toxicity tests should be collected, three aspects may be considered:

1) when effluent poses greatest potential for adverse effect;
2) when biological monitoring is conducted to look at linkages with effects in the exposure area; and,
3) season (in order to assess seasonal variability).

7.1.4 Sampling Locations

The effluent sample for sublethal toxicity testing will be collected at the mine's final discharge point that has potentially the most adverse impact on the environment, taking into account:

- monthly mass loading of the deleterious substances (as calculated in subsection 20(2) of the MMER); and,
- the manner in which the effluent mixes in the exposure area.

The mine should consider historical chemical characterization or sublethal toxicity data. In cases where it is not clear which discharge source has the greatest potential to impact the environment, mines may wish to use a single concentration sublethal toxicity test to determine the source with the greatest sublethal response.

To estimate the potency of the response from each discharge source, the "time to response" is observed and calculated as the test endpoint (e.g., *Ceriodaphnia* adults are exposed to undiluted effluent samples from each different effluent discharge and observations are made to how long it takes to see a 25% or 50% response). The test endpoint would be an LT25 (time to 25% mortality) or LT50 (time to 50% mortality) if survival was the key observation. The single concentration test would be the more cost effective approach to screening effluent sources to determine the discharge point with the greatest potential to impact the receiving environment.

7.1.5 Reporting

The results of sublethal toxicity testing will be submitted to the authorization officer as part of the Effluent and Water Quality Monitoring Report. This report will be submitted no later than March 31 of each year, for the previous calendar year's results, and will include, for sublethal toxicity testing, the following:

- dates when the samples were collected for sublethal toxicity testing;
- the location of the final effluent discharge point from which samples were collected for sublethal toxicity testing and data on how this point was chosen;
- the results of sublethal toxicity testing;
- a description of the QA/QC measures that were implemented, and the data related to the implementation of those measures.

7.2 Description of Freshwater and Marine Sublethal Toxicity Tests

The following are short descriptions for the freshwater and marine sublethal toxicity tests included in the EEM program.

For freshwater tests, laboratory or site water can be used as dilution/control water for the test. For marine or estuarine environments, the mine has a choice of using uncontaminated sea water or artificial sea water produced from hyper-saline brine (HSB).

For QA/QC, detailed records of all aspects of the samples, test organisms, culture maintenance, test conditions, equipment and test results are validated and kept by the laboratory. A reference toxicant test is used to establish the validity of effluent toxicity data. Successive reference toxicant data are plotted on a control chart. If results are within expected limits, the performance of the batch of test organisms is ensured. Minimum level of reporting is outlined in the test method.

Test methodologies may include:

- an incubation chamber/room with controlled temperature and illumination;
- toxicity testing equipment (exposure chambers);
- algal, macrophyte, invertebrate and fish culture facilities;
- particle counter;
- holding aquaria;
- clean, non-chlorinated water or uncontaminated sea water; and,
- deionized/distilled water.

Technical personnel should be skilled in algae, macrophyte, invertebrate and fish culture and toxicity tests following aseptic techniques.

For more detailed descriptions, please refer to the specific test method documents.

7.2.1 Macrophyte Growth Inhibition Test Using *Lemna minor*

Purpose: Evaluate effects of effluent exposure on the growth of a freshwater plant. Result is expressed as the concentration where frond number or growth is reduced by 25% (statistical endpoint is an IC25).

Description: *Lemna minor* (also known as lesser duckweed or common duckweed) is a small vascular, macrophyte plant, which is found at or just below the surface in freshwater (ponds, lakes, stagnant waters and quiet areas of streams and rivers). It is a common macrophyte with nearly worldwide distribution from tropical to temperate zones. The freshwater duckweed plant grows in most regions of Canada. Its growth is rapid and occurs by lateral branching. Macrophytes are higher-order plants with flowers and are an essential component of ecosystems due to their importance as a food substrate and protective habitat for young larval fish.

Seven to 10 day old, rapidly growing plants (typical size of *Lemna minor* frond is 1 cm) are exposed to a series of concentrations of the effluent sample diluted with growth medium for seven days. During the test, the plants are incubated at 25 °C under continuous light and static conditions. Plants are acclimated to the test media for 24 h before testing. The leaves are counted and weighed at the end of the test and the growth is compared statistically to that of the controls. For a valid test, the number of leaves on control plants will increase by 10-fold at the end of the test.
The test requires approximately 1-2 L of effluent.

Methodology: Biological Test Method: Test for Measuring the Inhibition of Growth using the Freshwater Macrophyte, *Lemna minor* (Reference Method EPS 1/RM/37), March 1999, published by the Department of Environment.

Cost: $760 (average from 5 CAEAL/MEF accredited labs – Environment Canada March 1999 survey).

7.2.2 Algal Growth Inhibition Test Using *Selenastrum capricornutum*

Purpose: Evaluate effects of effluent exposure on the growth of a unicellular freshwater alga. Result is expressed as the concentration where the number of cells is reduced by 25% (statistical endpoint is an IC25).

Description: *Selenastrum capricornutum* is a non-motile, unicellular, crescent-shaped (40 to 60 μm^3) green alga that is found in most fresh waters in North America. Its uniform shape makes it ideal for enumeration with an electronic particle counter. Clumping seldom occurs since *Selenastrum* is free of complex structures and does not form chains. Growth is sufficiently rapid to accurately count cell numbers after 72h.

Axenic (i.e., aseptically prepared stock cultures containing only the test species) exponentially growing *Selenastrum* are exposed to the test solutions in a static, 96-well microplate. The algae are exposed to a dilution series of filtered effluent sample, over several generations, under constant temperature (24 °C) with continuous light for 72 h. The number of algal cells in the test concentrations is compared with the number in the control solutions. An effluent is considered toxic when a statistically significant, dose-dependent inhibition of algal growth occurs. The test requires <1L of effluent.

Methodology: Biological Test Method: Growth Inhibition Test using the Freshwater Algae, *Selenastrum capricornutum* (Reference Method EPS 1/RM/25), November 1992, amended in November 1997, published by the Department of Environment.

Or: Détermination de l'inhibition de la croissance chez l'algue *Selenastrum capricornutum* (Reference Method MA 500-S.cap.2.0), September 1997, published by the Centre d'expertise en analyse environnementale du Québec, Ministère de l'Environnement du Québec (MEQ).

Cost: $497 (average from 10 CAEAL/MEF accredited labs – Environment Canada March 1999 survey).

7.2.3 *Ceriodaphnia dubia* Reproduction and Survival Test

Purpose: Evaluate effects of effluent exposure on the reproduction of an invertebrate. Results are expressed as the concentration where the average number of young per female is reduced by 25% (IC25). If mortality is significant, it may be possible to calculate the lethal concentration for 50% of the test population (statistical endpoint is a LC50).

Description: *Ceriodaphnia* is a species of zooplankton that is abundant in lakes, ponds, and quiescent sections of streams and rivers throughout North America. In the test, *Ceriodaphnia* are separated so that there is one female adult animal per test vessel and 10 replicates per concentration. Young ceriodaphnids, less than 24 h old, are exposed to a minimum of five effluent concentrations and a control, at 20 °C. The test is completed when at least 60% of the surviving control organisms have had three broods of neonates (7 to 8 days). During each day of the test, adult survivorship is assessed, all young produced are removed and counted, and the test solutions are renewed. At the end of the test, the number of surviving adults and the number of young produced per adult are compared statistically between exposure concentrations and the controls. The test requires 3-4 L of effluent.

Methodology: Biological Test Method: Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia* (Reference Method EPS 1/RM/21), February 1992, amended in November 1997, published by the Department of Environment.

Cost: $947 (average from 11 CAEAL/MEF accredited labs – Environment Canada March 1999 survey).

7.2.4 Fathead Minnow *(Pimephales promelas)* Growth and Survival Test

Purpose: Evaluate effects of effluent exposure to an early life stage of fish. Results are expressed as the concentration where larval growth/survival is reduced by 25% (IC25). If mortality is significant, it may be possible to calculate the lethal concentration for 50% of the test population (statistical endpoint is a LC50).

Description: Fathead minnows are small, warm water fish found across North America in ponds and slow moving water areas of rivers. The female lays her eggs on the underside of hard surfaces where the male cares for the eggs until hatching. Fathead minnow larvae, less than 24 hours old, are exposed to a minimum of 5 concentrations of the effluent sample and a control for seven days at 20 °C. During the seven days, the larvae are fed brine shrimp twice per day and the test/control solutions are replaced once each day. At the end of the seven days exposure, the surviving larvae are counted and individual weights are measured to calculate the growth changes which are compared statistically between exposure concentrations and the controls. The test requires approximately 40L of effluent.

Methodology: Biological Test Method: Test of Larval Growth and Survival Using the Fathead Minnows (Reference Method EPS 1/RM/22), February 1992, amended in November 1997, published by the Department of Environment.

The test organism importation methodology (Environment Canada 1999b) should be referred to, where test organisms are purchased for immediate use in sublethal toxicity tests.

Cost: $1019 (average from 11 CAEAL/MEF accredited labs – Environment Canada March 1999 survey).

7.2.5 Rainbow Trout *(Oncorhynchus mykiss)* Embryo Development Test

Purpose: Evaluate effects of effluent exposure to an early life stage of fish. Results are expressed as the concentration where embryo viability is reduced by 25% (statistical endpoint is an EC25).

Description: Rainbow trout are common in clean, cold water streams in North America. In some areas, they are not a native species but have been introduced to the watershed. Rainbow trout are cultured throughout the country by commercial hatcheries. Adult rainbow trout migrate to shallow water to spawn in clean gravel. They bury their eggs in the rocks and gravel where the young live until the yolk sac is absorbed.

The embryo development test involves exposing recently fertilized rainbow trout eggs to a series of concentrations of the effluent sample for seven days at 14 °C. Test exposure solutions are renewed every day. Dead embryos are counted and removed during the test. At the end of the test, the embryo's viability is assessed and number of healthy embryos are counted for statistical comparison between test concentrations and the control. The test requires approximately 80-90 L of effluent.

Methodology: Biological Test Method: Toxicity Tests Using Early Life Stages of Salmonid Fish (Rainbow Trout) (Reference Method EPS 1/RM/28), July 1998, published by the Department of Environment.

Cost: $1787 (average from 6 CAEAL/MEF accredited labs – Environment Canada March 1999 survey).

7.2.6 Fertilization Assay Using Echinoids (Sea Urchins and Sand Dollars)

Purpose: Evaluate effects of effluent exposure on egg fertilization success of echinoids. Results are expressed as the concentration where the fertilized egg number is reduced by 25% (statistical endpoint is the IC25).

Description: Echinoids are considered to be structurally advanced and complex invertebrates. Seven species of sea urchins and three species of sand dollars are commonly found in the coastal marine waters of Canada. Mature and gravid male and female echinoids are stimulated to spawn by injecting potassium chloride. Semen from at least 3 males is pooled and numbers are adjusted to the desired sperm: egg ratio. Eggs from at least 3 females are pooled and numbers are adjusted to 2000eggs/mL. Sperm is exposed for 10 to 20 min (depending on the test option chosen) to a series of concentrations of the effluent sample. Eggs are then added to the test vessels for a 10 to 60 min additional exposure. Adding formalin terminates the test. Preserved eggs are counted (in the range of 100 to 200 eggs) and classified as either fertilized or not fertilized, under a microscope at 100X magnification.

For a valid test, the fertilization rate in the controls will be >50% but <100% and a positive and logical dose effect curve will be obtained. The test requires approximately 1 L of effluent.

Methodology: Biological Test Method: Fertilization Assay Using Echinoids (Sea Urchins and Sand Dollars) (Reference Method EPS 1/RM/27), December 1992, amended in November 1997, published by the Department of Environment.

It is recommended that Environment Canada (2001) guidance document on salinity adjustment be followed when adjusting the salinity of the effluent, dilution water and the preparation of the hypersaline brine (HSB).

Cost: $1140 (Toxicology Expert Working Group Report – Review of Pulp and Paper Cycle 1 and Recommendations for Cycle 2, May 1997).

7.2.7 Larval Growth and Survival Assay Using Inland Silverside (*Menidia beryllina*)

Purpose: Evaluate effects of effluent exposure on fish larvae. Result is expressed as the concentration where larval growth is reduced by 25% (IC25). If mortality is significant, it may be possible to calculate the lethal concentration for 50% of the test population (statistical endpoint is a LC50).

Description: The inland silverside is a small fish which populates a variety of habitats and tolerates a wide range of temperature (2.9 to 32.5 °C) and salinity (0 to 58 g/kg). Inland silverside is a multiple spawner, and spawning can be induced by diurnal interruption in the circulation of water in the culture tanks. The eggs adhere to vegetation in the wild or to filter floss in laboratory culture tanks. The larvae hatch in 6 to 7 days when incubated at 25 °C and maintained in seawater ranging from 5 to 30 g/kg.

Seven to 11-day old larvae are exposed to a minimum of 5 concentrations of the effluent sample and a control for 7 days in a static renewal system at 25 °C. During the seven days, the larvae are fed brine shrimp twice per day and the solutions are replaced once each day. At the end of the seven days exposure, the surviving larvae are counted and individual weights are measured to calculate the growth changes which are compared statistically between exposure concentrations and the controls. For valid test, average dry weight of control larvae will be ≥ 0.50 mg and control survival will be ≥ 80%. The test requires approximately 40 L of effluent.

Methodology: Short Term Methods for Estimating Chronic Toxicity of Effluent and Receiving Waters to Marine and Estuarine Organisms (Third Edition) (Reference Method EPA/600/4-91/003), August 1994, published by the U.S. Environmental Protection Agency.

It is recommended that Environment Canada (2001) guidance document on salinity adjustment be followed when adjusting the salinity of the effluent, dilution water and the preparation of the hypersaline brine (HSB).

The test organism importation methodology (Environment Canada 1999b) should be referred to, where test organisms are purchased for immediate use in sublethal toxicity tests.

* For the USEPA method, the minimum reporting outlined in Environment Canada test methodologies should be followed.

Cost: $1510 (Toxicology Expert Working Group Report – Review of Pulp and Paper Cycle 1 and Recommendations for Cycle 2, May 1997).

7.2.8 Larval Growth and Survival Assay Using Topsmelt (*Atherinops affinis*)

Purpose: Evaluate effects of effluent exposure on fish larvae. Result is expressed as the concentration where larval growth is reduced by 25% (IC25). If mortality is significant, it may be possible to calculate the lethal concentration for 50% of the test population (statistical endpoint is a LC50).

Description: Topsmelt are small fish which occur from the Gulf of California to Vancouver Island. Topsmelt reproduce from May through August, depositing eggs on benthic algae in the upper ends of estuaries and bays. Off-season spawning has been successful in a laboratory-held population. Spawning is induced by a combination of three environmental cues: lighting, tidal cycle, and temperature.

Nine to 15-day old larvae are exposed to a minimum of 5 concentrations of the effluent sample and a control for 7 days in a static renewal system at 20 °C. During the seven days, the larvae are fed brine shrimp twice per day and the solutions are replaced once each day. At the end of the seven days exposure, the surviving larvae are counted and individual weights are measured to calculate the growth changes which are compared statistically between exposure concentrations and the controls. For valid test, average dry weight of control larvae will be ≥ 0.85 mg and control survival will be $\geq 80\%$. The test requires approximately 40 L of effluent.

Methodology: Short Term Methods for Estimating Chronic Toxicity of Effluent and Receiving Waters to West Coast Marine and Estuarine Organisms (First Edition) (Reference Method EPA/600/R-95-136), August 1995, published by the U.S. Environmental Protection Agency.

It is recommended that Environment Canada (2001) guidance document on salinity adjustment be followed when adjusting the salinity of the effluent, dilution water and the preparation of the hypersaline brine (HSB).

The test organism importation methodology (Environment Canada 1999b) should be referred to, where test organisms are purchased for immediate use in sublethal toxicity tests.

* For the USEPA method, the minimum reporting outlined in Environment Canada test methodologies should be followed.

Cost: $1600 (Michelle Fennell, Pacific Environmental Science Centre, Personal Communication).

7.2.9 Sexual Reproduction Assay Using the Red Macroalga (*Champia parvula)*

Purpose: Evaluate effects of effluent exposure on the sexual reproduction of a marine red macroalga. Result is expressed as the concentration where the number of cystocarps is reduced by 25% (statistical endpoint is IC25).

Description: Mature plant body of *Champia parvula* is hollow, septate and highly branched. New cultures can be propagated asexually from excised branches, making it possible to maintain clonal material indefinitely. Two sexually mature male and 5 female branches of *Champia parvula* are exposed in a static system for 2 days to a series of concentrations of the effluent sample, followed by a 5 to 7 day recovery period in control medium. The recovery period allows time for the development of cystocarps on the female branches resulting from fertilization during the exposure period.

For a valid test, the female control mortality will be ≤ 20%, and the average number of cystocarps per female control plants will must be ≥ 10. The test requires approximately 2 L of effluent.

Methodology: Short Term Methods for Estimating Chronic Toxicity of Effluent and Receiving Waters to Marine and Estuarine Organisms (Third Edition) (Reference Method EPA/600/4-91/003), August 1994, published by the U.S. Environmental Protection Agency.

It is recommended that Environment Canada (2001) guidance document on salinity adjustment be followed when adjusting the salinity of the effluent, dilution water and the preparation of the hypersaline brine (HSB).

* For the USEPA method, the minimum reporting outlined in Environment Canada test methodologies should be followed.

Cost: $1760 (Toxicology Expert Working Group Report – Review of Pulp and Paper Cycle 1 and Recommendations for Cycle 2, May 1997).

7.3 Dilution Water Issues in Freshwater Sublethal Toxicity Testing

7.3.1 Dilution Water Selection

Sublethal toxicity testing methods to be followed during the mining sector EEM program define the conditions and procedures to be followed during culturing and testing (Environment Canada 1992a,b,c, 1998, 1999a; USEPA 1994a, 1994b, 1995). Some testing decisions are left to the discretion of the individual laboratories, as long as the standard test acceptability criteria can be achieved. For example, standard methods for testing *Ceriodaphnia*, fathead minnows, *Selenastrum* and rainbow trout allow the use of uncontaminated ground or surface water, dechlorinated tap water or reconstituted water as a source for the culture or the test control/dilution water as long as the water of choice supports a healthy culture and provides a valid test result (Table 7.2).

Most laboratories in Canada use a "standard laboratory" water for routine culturing and testing requirements. This water is generally supplied to the laboratory through the local municipality and is either deionized and reconstituted or dechlorinated, or a mixture of both. Advantages of using laboratory water include:

- it can be maintained at a consistent quality with minimal risk of contamination by undesirable and/or harmful chemicals or biota;
- regular monitoring of water chemistry and culture health, as well as reference toxicant testing, ensure that the water is of acceptable quality for toxicity testing;
- since cultures are maintained in laboratory water, no additional acclimation is needed for testing effluents or chemicals, when laboratory water is used as the control/dilution water; and,
- laboratory water is normally used in regulatory testing across Canada since it provides a measure of the inherent toxicity of the effluents and allows comparison of effluent quality over time.

During the mining EEM program, most sublethal toxicity tests will likely be performed using laboratory water as control/dilution water, in order that the results may be comparable among different laboratories, and over time. It is also likely that in many situations where there is measurable effluent toxicity that this toxicity will be attributed to inorganic substances such as metals and ammonia, whose potency will be affected by site-specific characteristics such as pH, alkalinity and hardness. These characteristics can be reproduced in the laboratory for cases in which test results, reflective of the site conditions, are desired. However, a mine may decide to test its effluent using unexposed surface water (as control and dilution water), providing the sample is collected upstream from the discharge or adjacent to the source but removed from it so the sample is not exposed to effluent. Alternatively, a reference area of similar physical/chemical characteristics from a mine site could be used to supply control/dilution water.

The use of unexposed surface waters can be especially helpful in obtaining the following information:

To estimate the mitigating or stimulatory effect of unexposed surface site water as dilution water on the expression of toxicity from the effluent discharge or effluent exposed surface water.

Although parallel testing of effluents and effluent exposed surface waters using site water and hardness-adjusted laboratory water have produced similar results (BEAK 1998, 1999), it is impossible to simulate all the physical/chemical characteristics of site water using laboratory water. Therefore, if characteristics of site water, other than hardness, alkalinity and pH, are major influences in the expression of toxicity, it may be necessary to test using site water in order to understand site specific effects.

Unexposed surface water includes mine site collected water that has been collected upstream of a mine effluent discharge or from a nearby reference area. Unexposed surface water from the mine site area may vary in physical, chemical and biological characteristics over time.

Disadvantages of using unexposed surface water as dilution/control water include:

- relatively large volumes of unexposed surface water may be needed for testing, therefore additional expense is incurred for the collection, shipment and storage of site water samples (refer to Section 7.2.1 to 7.2.9 for volumes needed);
- some laboratory organisms will need acclimation to the unexposed surface water if it is significantly different in physical/chemical characteristics from the laboratory water (refer to Section 7.3.3); and,
- mandatory screening of the water through 60µm mesh to ensure indigenous populations of micro or macro organisms present in surface water do not impair the health of laboratory test organisms.

In spite of its practical disadvantages as control/dilution water, unexposed surface water is still preferred in many cases. The advantages include:

- it reflects the physical/chemical characteristics of the receiving environment; and,
- tests conducted may better reflect the influence of receiving environment characteristics (e.g., pH, hardness) on toxicant potency than laboratory water .

It should be noted that when using site water, problems may arise with reference water exhibiting sublethal responses. Beak (1998) attributed this problem to indigenous populations of micro-organisms infecting the laboratory organisms. Beak staff found that boiling the water prior to use before testing was successful in reducing mortality.

Sprague (1997) prepared an extensive review of studies that compared toxicity test results to receiving water impact and concluded that effects measured in sublethal toxicity tests correlate with environmental effects most of the time, especially if water, collected upstream of the effluent discharge is used as the control/dilution water.

When the purpose of a sublethal toxicity test is to estimate site-specific effects of contaminants, unexposed surface water from the vicinity of the mine site is recommended for use as control/dilution by Environment Canada, the U.S. EPA and ASTM in their method and associated guidance documents (Environment Canada 1992a,b,c, 1998, 1999a; U.S. EPA 1994a; ASTM 1998). Site-specific toxicity may also be compared to that measured in a test using laboratory water to adjust national water quality criteria to better reflect individual sites (U.S.EPA 1994c).

Table 7-2: Specifications of the Environment Canada test methods and recommendations for collection, storage and use of site-collected dilution waters

Criteria	*Ceriodaphnia dubia*	Fathead minnow	*Selenastrum capricornutum*	*Lemna minor*	Rainbow Trout embryo
Acceptable Dilution Water					
For Culturing	Uncontaminated ground, surface, dechlorinated, reconstituted water, dilute mineral water, or receiving water	uncontaminated ground, surface, or dechlorinated water	growth medium	Hoagland's E+ medium	ground, surface, reconstituted, dechlorinated, receiving water
For Testing	Reconstituted, dechlorinated, uncontaminated ground or surface water, receiving water	reconstituted, dechlorinated, uncontaminated ground or surface water, receiving water	reagent water, uncontaminated receiving water, groundwater, surface water, or reconstituted water	modified APHA growth medium, SIS growth medium, receiving water	ground, surface, reconstituted, dechlorinated, receiving water
Site Water					
Collection Point	Upstream or adjacent from source but removed from effluent exposure	upstream or adjacent from source but removed from effluent exposure	upstream or adjacent from source but removed from effluent exposure	upstream or adjacent to source but removed from effluent exposure	upstream or adjacent to source but removed from effluent exposure
Collection Procedure	as for effluent	as for effluent	as for effluent	as for effluent	as for effluent
Acclimation Procedure	Recommends acclimating at least 2 generations of brood organisms before collecting neonates for tests	recommends acclimation of breeding stock prior to testing	none	recommends placing plants in site collected dilution water; 18 to 24 hours prior to testing	none
Acclimation Rationale	Recommended hardness ±20% of culture water range, alkalinity range ±20% of culture water	when outside hardness, alkalinity range ±20% of culture water recommended	n/a	need to be done for all types of test medium	n/a
Treatments	Recommend filter 60µm, boiling if necessary	recommend filter 60µm, boiling if necessary	filter 0.45µm	filter 1µm, then filter 0.22µm, nutrient-spiked	recommend 60µm filter or boiling if necessary
Storage	Preferably no more than 14 days; maximum of one month at 4ºC with no headspace	preferably no more than 14 days; maximum of one month at 4ºC with no headspace	preferably no more than 14 days; maximum of one month at 4ºC with no headspace	preferably no more than 14 days; maximum of one month at 4ºC with no headspace	preferably no more than 14 days; maximum of one month at 4ºC with no headspace
During Toxicity Testing	Include lab water control. If screening test shows impairment, treat water by boiling. If impairment remains, use hardness adjusted lab water.	Include lab water control. If screening test shows impairment, treat water by boiling. If impairment remains, use hardness adjusted lab water.	include reagent water control	include lab water control	include lab water control. If screening test shows impairment, treat water by boiling. If impairment remains, use hardness adjusted lab water.

The widespread acceptance of unexposed surface dilution water for predicting site-specific effects is based on knowledge regarding the interaction of contaminants with water quality characteristics. For example, metal toxicity is well known to be influenced by the physical/chemical characteristics of water, such as pH, alkalinity, hardness (reviewed by Wang 1986 and Sprague 1995). However, recent studies comparing results of toxicity tests on effluents and effluent exposed surface waters from four different mine sites indicated that a similar estimation of toxicity could be obtained from tests using unexposed surface water or laboratory water for dilution, especially if the laboratory water was adjusted to the hardness, alkalinity and pH of the site water (BEAK 1998, 1999). This finding indicates that the use of site-collected dilution water may not always be necessary since laboratory waters can be prepared to reflect characteristics of the site water such as hardness, pH and alkalinity.

One important aspect of some site waters known to mitigate the effects of metal toxicity, by reducing metal bioavailability, is naturally elevated levels of organic carbon. Ranges of dissolved organic carbon (DOC) have been measured as high as 58 mg/L in some Ontario lakes (Nearly 1990, as cited in Welsh et al. 1993) and such waters may influence the expression of metal toxicity in effluent or effluent exposed surface water matrices. However, in parallel toxicity tests of mine effluent conducted using site water of moderately high organic carbon (total organic carbon of 9.4 mg/L) and hardness-adjusted laboratory water showed no significant difference in organism response (BEAK 1999). Of note is that, high organic carbon content may not always result in reduced metal bioavailability. In waters of high hardness, concentrations of calcium and magnesium may be high enough to bind with humic acid (which make up the majority of organic carbon). Therefore, humic acid binding sites would be limited and unavailable to tie up free metal ions. In this situation, decreases in effluent sublethal toxicity due to high receiving water DOC would not occur (Winner 1985).

To compare toxicity of the discharge or effluent exposed surface water relative to an impaired upstream water.

If upstream water is contaminated by non-point or upstream point sources of pollution that are unrelated to the mine operation, a mine may decide to use that water for test dilution purposes in toxicity testing in order to provide an appropriate comparison of test organism responses, as long as the upstream water can support the health of the test organisms. If the upstream water cannot support health of the test organisms, it could be tested separately in a dilution series to quantify its effects with uncontaminated reference site or regular laboratory water used for test control and dilutions.

7.3.2 Collection, Shipment and Storage of Samples for Sublethal Toxicity Testing

Collection and Shipment

Methodologies are similar for the collection and shipment of site-collected dilution water, effluent or effluent exposed surface water. It is important that containers used for sampling, transportation, and storage of dilution water, effluents and/or effluent exposed surface water are made of non-toxic material, preferably plastic containers made of polyethylene or polypropylene. Containers used for sampling need to be rinsed with the waters which are to be tested and filled completely to eliminate any air space. All containers will be labeled with the sample name and date, then properly sealed and fastened with tape. According to the Environment Canada standard methods, samples which are >7°C upon sampling should be cooled to 1-7°C prior to shipping. As necessary, gel packs or other means of refrigeration should be used to assure that sample temperature remains within 1-7°C during transit (Environment Canada 1992 a,b,c, 1998, 1999a). Samples cannot be allowed to freeze during transport. A recommended approach is to use plastic 10 L cubitainers for the sample, placed inside a cooler and chilled with crushed ice or large ice packs to keep samples cold during transport. The cooler should be sealed with strong packing tape and initialed in several places, to ensure the integrity of the sample. It is preferable to ship the sample through a reliable courier which will deliver within 24 hours but the sample has to be delivered within 72 hours of sampling. A sample submission form/chain of custody will be filled out and included in all shipments.

Table 7.3 provides estimates of the volumes of site water needed for performing a suite of mining EEM tests, and includes estimates for effluents or effluent exposed surface water volume. As recommended by the U.S. EPA (1994a) site-collected dilution water samples should be representative of the water body, unaffected by recent run-off or erosion events that may cause the water to have a higher total suspended solids (TSS) concentration.

Table 7-3: Dilution/control water and corresponding effluent volumes for sublethal toxicity tests*

Test	Dilution Water Volume (L)	Effluent Volume (L)
Ceriodaphnia dubia	10	4
Fathead minnow	45	20
Selenastrum capricornutum	1	1
Lemna minor	5 for static, 12 for static-renewal	2 for static, 5 for static-renewal
Rainbow Trout (embryo test)	300	125
Testing Series	**Dilution Water Volume (L)**	**Effluent Volume (L)**
Ceriodaphnia, Fathead, *Selenastrum*, *Lemna* (static)	65	30
Ceriodaphnia, Fathead, *Selenastrum*, *Lemna* (static-renewal)	72	35
Ceriodaphnia, Rainbow Trout embryo, *Selenastrum*, *Lemna* (static)	325	140
Ceriodaphnia, Rainbow Trout embryo, *Selenastrum*, *Lemna* (static-renewal)	330	150

All volumes are calculated assuming one control and 5 concentrations. Fathead minnow assumes 300mL test volume and 4 replicates, and *Lemna minor* assumes 150mL test volume and 4 replicates.

* estimated effluent volumes for marine/estuarine sublethal toxicity tests are outlined in Sections 7.2.6, 7.2.7, 7.2.8 and 7.2.9.

Storage

Uncomposited effluent and effluent exposed surface water samples should be stored, without headspace at 4°C for a maximum of three days prior to test initiation. If a mine's effluent has been unavoidably delayed but sample integrity was retained, sublethal toxicity test results from tests up to four days after collection may be considered valid, if the AO was notified and concurs with the sample delay. Separate containers are used for different days of testing in the static renewal tests, according to the most recent version of Environment Canada's test methods.

Site-collected dilution water sample should be used as soon as possible. The U.S. EPA states that water not be held longer than 96 hours prior to initiation of toxicity tests (U.S. EPA 1994a). However, it may not always be possible to initiate tests immediately, especially if screening tests are to be performed on the water (refer to Section 7.3.4). Environment Canada does not specify a maximum time period for holding site-collected dilution water and studies in 1999 have shown that uncontaminated water may be held for at least one month at 4°C, without headspace, with no significant reduction in water quality or organism response (BEAK 1999). However, unexposed water samples tested were collected in early spring when background bacteria concentrations were probably low. Site water samples collected during summer months may be problematic due to bacterial contamination brought on by higher surface water temperatures.

According to Environment Canada test methods, site-collected dilution water should be treated in the same manner as effluent or effluent exposed surface water when stored, implying that containers should not be composited (Environment Canada, 1992 a,b,c, 1998,1999a). However, it may be more advisable to composite a dilution water, especially if screening tests are to be performed on the water prior to initiation of the effluent (or effluent exposed surface water) toxicity test.

7.3.3 Test Organism Acclimation

Pre-acclimation of culture organisms is recommended prior to exposure to site water. Since the purpose of using site water for test control and dilutions is to more accurately predict receiving water impact, the most accurate prediction should be achieved using organisms adapted to the physical/chemical conditions of the receiving environment. Lloyd (1965) found that fish cultured in hard water need to lose calcium before they are as sensitive to metals as fish cultured in soft water. Therefore, if site water to be used for test dilution is softer than the lab water, pre-acclimating fish to the site water conditions would provide enough time for loss of calcium prior to test initiation. Alternatively, if site-collected water is higher in hardness, then pre-acclimation would allow fish species to accumulate calcium prior to testing.

The Environment Canada methods for *Ceriodaphnia dubia* and fathead minnows recommend that cultures be maintained in water of similar hardness, alkalinity and pH (i.e., within 20%) to the site water used for test dilution (Environment Canada 1992 a,b). Since many Canadian mines are located beside rivers or lakes of low hardness, it is likely that some acclimation of laboratory cultures would be necessary. Beak International Inc. (BEAK) developed a pre-acclimation procedure during the 1997 AETE study, later refined in the 1999 study, for fathead minnows and *Ceriodaphnia dubia* (BEAK 1997, 1999). Based on the expected hardness, alkalinity and pH of the site water, cultures are gradually introduced to laboratory water of decreasing hardness over several days until the appropriate hardness is reached. This procedure was adapted from that used by B.A.R. Environmental Inc. (BAR) during its 1996 AETE study, in which cultures were gradually acclimated to hardness-adjusted laboratory water and site water if screening of un-acclimated organisms showed impairment to laboratory organisms (BAR 1997).

For pre-acclimation of cultures, laboratory water of reduced hardness may be prepared by diluting standard laboratory water with deionized water. Hardness can be increased by adding salts, used for the preparation of reconstituted water in the appropriate amounts (Table 7.4). When hardness-adjusting water, it is important to keep the alkalinity level appropriate to the hardness since alkalinity affects the speciation of metals (U.S. EPA 1994c, Lauren and McDonald 1986). Appropriate hardness and alkalinity relationships are available in Table 7.4, and additional values may be interpolated (U.S. EPA 1994c).

Detailed procedures for pre-acclimation of *Ceriodaphnia dubia*, fathead minnows test methods are described below. References to hardness, assume a corresponding change in alkalinity and pH. No pre-acclimation procedures are described for the rainbow trout test, since eggs are delivered from the hatchery and used within 24 hours in testing (EC 1995). Similarly, *Selenastrum* and *Lemna* cultures are maintained in standard culture media that are different from standard testing media, although *Lemna* does allow for some pre-acclimation of cultures, since plants are transferred to the test medium 18 to 24 hours prior to initiation of testing (Environment Canada 1999a).

Table 7-4 Preparation of different hardness/alkalinities

Water Type	**Reagent Added (mg/L)**[1]			**Final Water Quality**			
	$NaHCO_3$	**$CaCSO_4$-$2H_20$**	**$MgSO_4$**	**KCL**	**pH**[2]	**Hardness**[3]	**Alkalinity**[3]
Very Soft	12.0	7.5	7.5	0.5	6.4 - 6.8	10 - 13	10 - 13
Soft	48.0	30.0	30.0	2.0	7.2 - 7.6	40 - 48	30 - 35
Moderately H	96.0	60.0	60.0	4.0	7.4 - 7.8	80 - 100	60 - 70
Hard	192.0	120.0	120.0	8.0	7.6 - 8.0	160 - 180	110 - 120
Very Hard	384.0	240.0	240.0	16.0	8.0 - 8.4	280 - 320	225 - 245

From U.S. EPA (1994a)

[1] Add reagent grade chemicals to deionized water.

[2] Approximate equilibrium pH after 24h of aeration.

[3] Expressed as mg $CaCO_3$/L.

Ceriodaphnia dubia

Ceriodaphnia cultures are initiated and maintained according to the Environment Canada standard method. To pre-acclimate cultures to the hardness of the site-collected dilution water, a brood of neonates, less than 24 hours old, is initiated in water, reduced in hardness by 20% from that of laboratory water, by addition of deionized water. Each day, organisms are transferred to new solutions, decreased by a further 20 to 30% in hardness. Once the desired hardness is reached (within approximately one week) and the culture organisms pass the health criteria of the Environment Canada method (i.e., production of at least three broods, total neonate production of at least 15 per adult, less than 20% adult mortality), a new brood is initiated. The second generation cultures are maintained in the hardness-adjusted water until organisms pass the health criteria (approximately one week). Selenium and vitamin B12 are added to the hardness-adjusted culture water, if low in hardness, as recommended by the standard method.

Fathead Minnows

Acclimation and pre-acclimation procedures carried out by BAR and BEAK during the 1996 and 1997 AETE studies used breeding tanks of fathead minnows, gradually changed in hardness and alkalinity to that of the applicable site water (Hardness Adjusted Laboratory Water or HALW) (BAR 1997, BEAK 1998). Eggs were collected from the adjusted-water tanks and reared at the HALW until hatching. However, preliminary work completed by BEAK in early 1999 showed that eggs could be hatched out in HALW without prior acclimation of the breeding tanks. By omitting breeding tanks from the pre-acclimation procedure, the volume of HALW needed for pre-acclimation is reduced while maintaining a supply of eggs for hatching in a number of different water types. Eggs are hatched according to the Environment Canada standard method, with fresh water renewal every 24 hours.

7.3.4 Screening Tests

Once organisms are pre-acclimated to the physical/chemical conditions of the site water, they may be exposed to the site water in screening tests. If organisms exposed to the site water meet the test method control acceptability criteria, the site water may be considered suitable for use as dilution water provided the site water meets the control validity criteria for the test method. Comparison of the site water response to the laboratory response in a screening test may reveal a statistically-significant reduction in growth, reproduction and/or survival, even though the site water exposed group is within the control acceptability criteria for the test. As long as the site water meets the control validity criteria for the test method, it may be considered suitable for testing.

7.3.4.1 Screening Test Impairment

If organisms are given an adequate opportunity to gradually acclimate to the physical/chemical characteristics of the site water (by exposure to hardness-adjusted laboratory water) then impairment observed during screening tests should not be due to shock exposure to different water quality characteristics such as hardness, alkalinity or pH. Therefore, impairment would likely be due to the presence of harmful biological agents or toxicants.

Special attention should be paid to any 100% site water exposures showing significant mortality in one or more replicates. Microbiological organisms, present in the site water, may impair the health of test organisms and anecdotal evidence from several ecotoxicity laboratories indicates that impairment by indigenous micro-organisms usually occurs after a few days exposure. For example, this has been manifested in fathead minnow tests conducted at BEAK, as a sudden onset of significant mortality in the site-water control, often in one or two replicates only. Occasionally, evidence of fungal or bacterial growth may be observed in the test vessels. If such contamination is indicated by a screening test, then acclimation of the organisms will most likely not result in a removal of impairment. Therefore, either the site water should be treated by a suitable means to remove the impairment (i.e., boiling or UV treatment, see below) or hardness-adjusted laboratory water should be used as a surrogate.

Limited laboratory trials of site water showed that impairment could be removed by boiling site water gently for 10 minutes and cooled prior to use in testing. Other treatments have been reported in the literature, such as UV light and 0.45 μm filtration (Grothe and Johnson 1996; Kszos et al. 1997). If site water is collected during late spring to early fall, then some form of biological contamination should be expected (unless experience with a particular site water indicates otherwise) and precautions such as boiling should be taken.

If the impairment is due to chemical contamination, the suitability of the site water for use as a dilution and control water is questionable, even though cultures may be acclimated to naturally high levels of metals in site water (see Section 7.3.3). If the cultures are exposed to higher levels of contaminants, post-acclimation can result in either higher or lower sensitivity of laboratory organisms, depending on the contaminant, organism and the water characteristics and the utility of the post screening test acclimation procedure becomes questionable. If impairment is detected in a screening test, the recommended procedure is to attempt treatment by boiling or to use hardness-adjusted laboratory water as a surrogate dilution water. If site water is suspected of being contaminated, including a boiled site water exposure in screening tests would resolve the question of the effectiveness of that treatment for the site water of interest.

7.3.5 Effluent or Effluent Exposed Surface Water Toxicity Tests

Table 7.2 summarizes the recommendations for collecting, storing and using site-collected dilution waters in sublethal toxicity testing.

Once site water has been deemed suitable for testing, toxicity tests may be initiated as per the appropriate Environment Canada testing method. In addition to the site water control, an additional control, using water from the laboratory culture, needs to be included in testing to serve as a check of culture health and site water quality. No additional control is necessary when the test control dilution water is the same as the culture water. In the case of the *Lemna minor* and *Selenastrum capricornutum* tests, the laboratory control would be the standard test growth medium, specified in the standard method.

If site-collected water is deemed unsuitable for use as test control and dilution water, a treatment such as boiling should be attempted to remove the impairment. If successful, the treated site-collected dilution water should be used in testing. However, boiling large volumes of water may not be practical for tests such as the rainbow trout test, and hardness-adjusted laboratory water may be a more suitable alternative. If no practical treatment can be found to remove impairment to test organisms caused by the site-collected dilution water, then hardness-adjusted water should be used as the dilution water with pre-acclimated organisms. Care should be taken to match the pH of the characteristics of the site-collected dilution water as closely as possible, since the expression of metal toxicity may change within the acceptable pH range of the test.

7.4 Quality Assurance and Quality Control (QA/QC)

In addition to the quality assurance outlined in the individual sublethal toxicity test methods, the following is also recommended: :

1) reference toxicant test be conducted in the same manner as the effluent or effluent exposed surface water test;
2) a reference toxicant test have been conducted within 30 days of the effluent or effluent exposed surface water test;
3) test-specific validity criteria be met in all effluent sublethal testing conducted;
4) sublethal toxicity testing be initiated within 3 days of sample collection;
5) quantitative endpoints be provided for all sublethal toxicity tests conducted on effluent or effluent exposed samples; and
6) test endpoints between 0.1% and 100% be bracketed by at least one test concentration.

As well, it is highly recommended that laboratories contracted by the mining industry to conduct sublethal toxicity testing should be accredited by the Standard Council of Canada through the Canadian Association for Environmental Analytical Laboratories (CAEAL), the Ministère de l'Environnement du Québec or an organization offering an equivalent level of accreditation. These accreditation organizations follow International Standards Organization (ISO) standards, which focus on the quality system and practices of a laboratory. This level of quality assurance should be a prerequisite for selecting a testing laboratory.

7.5 Use of Sublethal Toxicity Testing in Resolving Confounding Influences

Sublethal toxicology data also has the potential utility to aid in the resolution of confounding factors. The Toxicology Subgroup recognized this as a third use for sublethal toxicity data.

To estimate the relative contribution of mine effluent releases and other natural and/or anthropogenic influences on sublethal toxicity in the same receiving water body

Sublethal toxicity test data can be used to deal with situations where there are confounding influences which make the cause of receiving environment effects difficult to interpret. The choice of where to use sublethal toxicity tests in this application is up to the mine operator and the nature of the confounding influences situation. However, the confounding influence scenario where sublethal toxicity testing would be most relevant is the multiple point source discharge and/or non-point source input situation. There is no particular set of sublethal tests or frequency of monitoring recommended for this use due to the site-specific nature of confounding influence situations.

Estimates of sublethal toxicity can help in understanding the relative contribution of diverse industrial or municipal discharges to effects on aquatic organisms in the receiving-water, whether the discharges are from up-stream point or non-point sources (e.g., acid rock drainage, seepage form an abandoned mine, municipal landfill leachate, etc.) or the mine's property. The upstream contribution to an the observed environmental effect can be estimated, given surface water sublethal toxicity data, discharge flow, and features of dispersion into the receiving environment. If plumes from different discharges at a mine site overlap, more effort is necessary to distinguish the toxic contributions of the mine's discharge sources vs. up-stream sources. Samples of surface water from key locations in the near-field exposure areas could be tested, to estimate the combined toxic contribution of the sources.

A three step procedure for sorting out the contribution from different sources of sublethal toxicity to the near field receiving environment is outlined below:

- Conduct a battery of sublethal toxicity tests on samples collected from all significant discharge sources from the mine's property. Use standard laboratory water for test dilutions and control, or unexposed site water. This estimates the absolute sublethal toxicity of each mine site discharge. Repeat the sampling and testing on any discharge which is known to be variable in toxicity, to obtain an estimate of the degree of variability.

- Conduct a parallel battery of sublethal toxicity tests for each discharge to a river, using water collected directly upstream from the point of discharge for dilution and control. For lakes or estuaries, carry out the parallel battery of tests by collecting control/dilution water from outside the zone immediately affected by the discharges. Separate and simultaneous controls should be run using standard uncontaminated water as a quality assurance measure. It should be recognized that "upstream" sources of control/dilution water might already be contaminated by other effluent discharges or sources of toxicants. Accordingly, the upstream dilution water might contribute to significant effects on growth or reproduction, in concentrations of the effluents being studied, or even in control vessels. This would not invalidate the results since the purpose of the investigation is to evaluate the relative contributions of discharges to the total toxicity of the receiving water.

- Confirmation of the relative contribution of discharges is recommended and can be achieved by conducting sublethal toxicity tests on samples of surface water from the water body receiving the discharges (so-called "ambient" tests). This can aid in:
  a) confirming whether or not an effluent has a measurable toxicity after mixing into the receiving water;
  b) estimating the persistence in the receiving water, of toxicity from all contributing sources; and
  c) determining the combined toxicity resulting from the mixing of all point and non-point sources, as an estimate of the overall effect on the receiving environment.

Testing samples of surface water which receives discharges or toxicants from multiple sources should be done synoptically and ideally during low-flow or worst-case periods. At a minimum, sampling should be carried out over as short a period of time as possible (e.g., 1 or 2 days). Repeated rounds of sampling and testing would be desirable if the toxicity of the discharges were variable. The above guidance on conductance of toxicity assessment studies to estimate the contribution of multiple-discharge sources to instream effects is based on the eight site investigations conducted under the U.S. EPA Complex Effluent Toxicity Testing Program (CETTP). Detailed reports on these studies were prepared by Mount and Norberg-King (1985, 1986), Mount et al. (1984, 1985, 1986a, 1986b, 1986c) and Norberg-King and Mount (1986). The results of CETTP program, including independent critiques and re-analyses, were reviewed by Sprague (1997) during a project commissioned by the Aquatic Effects Technology Evaluation (AETE) Program. Sprague concluded that the U. S. EPA CETTP studies provided valid findings which should be considered by Canadian metal mining companies when designing aquatic environmental monitoring programs at their mine sites.

Mines may elect to conduct additional investigations where the most sensitive species in the effluent produced an IC25 of less than 30%. The following are additional recommended investigations. At a minimum, the most sensitive triggered test species can be used to estimate the geographic extent of the potential response. Alternatively, the results of the triggered toxicity test(s) may lead to other recommended laboratory or field monitoring tools.

A tiered approach to resolving the confounding influences is recommended starting with:
1) re-testing with the sublethal test which provided the most sensitive IC25 result using upstream or reference site water for test control and dilution's; or,
2) receiving water toxicity testing with samples collected from the area where a sublethal response is predicted.

7.5.1 Microtox Chronic Test

The Toxicology Subgroup recommended that the Microtox Chronic test could be used for screening effluents or effluent exposed surface waters. BAR (1997) found that the

Microtox chronic test had poor sensitivity. However, more recent studies comparing the Microtox chronic test to the *Selenastrum* growth inhibition test showed greater promise for the Microtox test. As a result, the Microtox chronic test could be used as a screening tool or in situations where there are multiple discharge points at a mine site. A brief summary of the Microtox Chronic test is provided below.

Purpose: Evaluate effects of effluent exposure on light production by the naturally luminescent marine bacteria, *Vibrio fischeri*. The result is expressed as the concentration where light output is reduced by 25% (statistical endpoint is an IC25).

Description: The chronic test is a development of the acute test, using the same species and the same incubator/test system. The Microtox chronic test is described in the manufacturer's handbook "Measuring Chronic Toxicity Using Luminescent Bacteria" (Microbics Corp. 1994).

A bacterial culture is incubated with concentrations of the effluent at 27 °C. After 22 hours, the amount of light output of the bacteria is measured.

For QA/QC, detailed records of all aspects of the samples, test organisms, test conditions, equipment and test results are validated and kept by the laboratory. A reference toxicant is used to establish the validity of effluent toxicity data. Successive reference toxicant data are plotted on a control chart. If results are within expected limits, the performance of the batch of test organisms is ensured.

Cost: Approximately $200 per effluent sample.

7.6 Investigation of Cause

If a mine wishes to investigate the cause of the sublethal toxicity response seen in their effluent, investigative procedures prepared by the U.S. Environmental Protection Agency are available to characterize and identify the toxic constituents of the effluent (U.S. EPA 1992). These procedures have been successfully used by some mines in Canada.

7.7 References

ASTM (American Society for Testing and Materials). 1998. Standard Guide for Conducting Acute Toxicity Tests on Aqueous Ambient Samples and Effluents with Fishes, Macroinvertebrates, and Amphibians, Designation E1192-97. Annual Book of ASTM Standards. Section 11: Water and Environment Technology.

BAR (B.A.R. Environmental Inc.). 1997. Toxicity Assessment of Mining Effluents using Upstream or Reference Site Waters and Test Organism Acclimation Techniques. Aquatic Effects Technology Evaluation Program. AETE Project 4.1.2a.

BEAK (Beak International Incorporated). 1999. Effects of Dilution Water on The Results of Sublethal Toxicity Tests. *In progress.*

BEAK (Beak International Incorporated). 1998. Additional Tool Evaluations. Aquatic Effects Technology Evaluation Program. Beak Reference 20776.1.

CCREM (Canadian Council of Resource and Environment Ministers. 1987. Canadian Water Quality Guidelines. Task Force on Water Quality Guidelines. Ottawa, Canada.

Environment Canada 1992a. Biological Test Method: Test of Larval Growth and Survival Using Fathead Minnows, Environmental Technology Centre, Ottawa, Ontario, Report EPS 1/RM/22, February 1992, Amended in November 1997.

Environment Canada 1992b. Biological Test Method: Fertilization Assay with Echinoids (Sea Urchins and Sand Dollars), Environmental Technology Centre, Ottawa, Ontario, Report EPS 1/RM/27, December 1992, Amended in November 1997.

Environment Canada 1992c. Biological Test Method: Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia*, Environmental Technology Centre, Ottawa, Ontario, Report EPS 1/RM/21, February 1992, Amended in November 1997.

Environment Canada 1992d. Biological Test Method: Growth Inhibition Test Using the Freshwater Alga *Selenastrum capricornutum*, Environmental Technology Centre, Ottawa, Ontario, Report EPS 1/RM/25, November 1992, Amended in November 1997.

Environment Canada 1998. Biological Test Method: Toxicity Tests Using the Early Life Stages of Salmonid Fish (Rainbow Trout), Environmental Technology Centre, Ottawa, Ontario, Report EPS 1/RM/28, Second edition, July 1998).

Environment Canada 1999a. Biological Test Method: Test for Measuring the Inhibition of Growth Using the Freshwater Macrophyte *Lemna minor*, Environmental Technology Centre, Ottawa, Ontario, Report EPS 1/RM/37, March 1999.

Environment Canada, 1999b. Recommended Procedure for the Importation of Test Organisms for Sublethal Toxicity Testing. Environmental Technology Centre, Ottawa, Canada.

Environment Canada, 2001. Revised procedures for Adjusting the Salinity of Effluent Samples for Marine Sublethal Toxicity Testing conducted under Environmental Effects Monitoring (EEM) Programs. Method Development and Application Section, Environmental Technology Centre, Ottawa, Canada.

ESG International Inc. (ESG). 1999. AETE Synthesis Report of Selected Technologies for Cost-Effective Environmental Monitoring of Mine Effluent Impacts in Canada, AETE Project 4.1.4, March 1999, CANMET, Natural Resources Canada, Ottawa, Ontario.

Grothe DR and Johnson DE. 1996. Bacterial Interference in Whole Effluent Toxicity Tests. Environmental Toxicology and Chemistry. 15(5): 761-764.

Howarth RS and Sprague JB. 1978. Copper Lethality to Rainbow Trout in Waters of Various Hardness and pH. Water Research. 12: 455-462.

Kszos LA, Stewart AJ and Sumner JR. 1997. Evidence that Variability in Ambient Fathead Minnow Short-term Chronic Tests is Due to Pathogenic Infection. Environmental Toxicology and Chemistry. 16(2): 351-356.

Microbics Corporation, Measuring Chronic Toxicity Using Luminescent Bacteria, 1994.

Ministère de l'Environnement du Québec (MEQ). 1997. Détermination de l'inhibition de la croissance chez l'algue *Selenastrum capricornutum* (Reference Method MA 500-S.cap.2.0), September 1997, published by the Centre d'expertise en analyse environnementale du Québec, Ministère de l'Environnement du Québec (MEQ).

Mount DI and Norberg-King TJ (eds). 1985. Validity of effluent and ambient toxicity tests for predicting biological impact, Scippo Creek, Circleville, Ohio. U.S. Environmental Protection Agency, Washington D.C. EPA 600/3-85/044.

Mount DI and Norberg-King TJ (eds). 1986. Validity of effluent and ambient toxicity tests for predicting biological impact, Kanawha River, Charleston, West Virginia. U.S. Environmental Protection Agency, Washington D.C. EPA 600/3-86/006.

Mount DI et al. 1984. Effluent and ambient toxicity testing and instream community response on the Ottawa River, Lima, Ohio. U.S. Environmental Protection Agency, Washington D.C. EPA 600/3-84/080.

Mount DI et al. 1985. Validity of effluent and ambient toxicity testing for predicting biological impact on Five Mile Creek, Birmingham, Alabama. U.S. Environmental Protection Agency, Washington D.C. EPA 600/8-85/015.

Mount. DI et al. 1986a. Validity of effluent and ambient toxicity tests for predicting biological impact, Naugatuck River, Waterbury, Connecticut. U.S. Environmental Protection Agency, Washington D.C. EPA 600/8-86/005.

Mount DI et al. 1986b. Validity of effluent and ambient toxicity tests for predicting biological impact, Back River, Baltimore Harbor, Maryland. U.S. Environmental Protection Agency, Washington D.C. EPA 600/8-86/001.

Mount DI et al. 1986c. Validity of ambient toxicity tests for predicting biological impact, Ohio River near Wheeling, West Virginia. U.S. Environmental Protection Agency, Washington D.C. EPA 600/3-85/071.

Norberg-King TJ and Mount DI (eds). 1986. Validity of effluent and ambient toxicity tests for predicting biological impact, Skeleton Creek, Enid, Oklahoma. U.S. Environmental Protection Agency, Washington D.C. EPA 600/8-86/002.

Sprague J. 1995. Factors that Modify Toxicity. In: *Fundamentals of Aquatic Toxicology*, 2nd Edition. Gary Rand (ed.) Taylor and Francis, Washington.

Sprague J. 1997. Review of methods for sublethal aquatic toxicity tests relevant to the Canadian metal-mining industry. Aquatic Effects Technology Evaluation Program. Natural Resources Canada. Canadian Centre for Mineral and Energy Technology.

U.S. EPA (U.S. Environmental Protection Agency), 1992. Toxicity Identification Evaluation: Characterization of Chronically Toxic Effluents, Phase1 . EPA 600/6-91/005F, Office of Research and Development.

U.S. EPA. 1994a. Short-term Methods for Estimating the Toxicity of Effluents and Receiving Water to Freshwater Organisms. Third Edition. Environmental Monitoring Systems Laboratory, U.S. Environmental Protection Agency, Cincinnati, OH 45268. EPA /600/4-90/027.

U.S. EPA 1994b. Short Term Methods for Estimating Chronic Toxicity of Effluent and Receiving Waters to Marine and Estuarine Organisms. Third Edition. Environmental Monitoring Systems Laboratory, U.S. Environmental Protection Agency, Cincinnati, OH 45628. EPA/600/4-91/003, August 1994.

U.S. EPA. 1994c. Interim guidance on determination and use of water-effect ratios for metals. EPA 823-B-94-001. Office of Water

U.S. EPA. 1995. Short term methods for estimating the toxicity of effluents and receiving water to West Coast organisms. First Edition, Environmental Monitoring Systems Laboratory, U.S. Environmental Protection Agency, Cincinnati, OH, EPA 600/R-95/136.

Wang W. 1997. Factors Affecting Metal Toxicity to (and Accumulation by) Aquatic Organisms - An Overview. Environment International 13: 437- 457.

Welsh PG, Skidmore JF, Spry DJ, Dixon DG, Hodson PV, Hickie BE. 1993. Effect of pH and Dissolved Organic Carbon on the Toxicity of Copper to Larval Fathead Minnows *(Pimephales promelas)* in Natural Lake Waters of Low Alkalinity. Canadian Journal of Fisheries and Aquatic Sciences. 50: 1356-1362.

Winner RW. 1985. Bioaccumulation and Toxicity of Copper as Affected by Interactions Between Humic Acid and Water Hardness. Water Research. 19: 449-455.

Table of contents

8. Sediment Monitoring 8-1

8.1 Collection of Sediment Samples 8-1
8.1.1 Field Measurements and Observations 8-2
8.1.2 Criteria for Selection of a Sample Collection Device 8-2
8.1.3 Collection Device Penetration Depth 8-8
8.1.4 Sample Size (volume of sediment) 8-9
8.1.5 Replicate Samples 8-9
8.1.6 Grab Sampler Operation 8-10
8.1.7 Criteria of Acceptability of Samples 8-11
8.1.8 Sediment Parameters that May be Measured in the Field 8-12

8.2 Sample Handling and Analysis 8-12
8.2.1 Procedures for Handling of Sediment Samples 8-12
8.2.2 Subsampling of Sediment Grab Samples 8-13
8.2.3 Compositing Sediment Grab Samples 8-13
8.2.4 Sample Containers 8-13
8.2.5 Transportation and Storage of Sediment Samples 8-14
8.2.6 Laboratory Test Sample Preparation 8-15
8.2.7 Prevention of Sediment Sample Contamination 8-16

8.3 Sediment Sample Analysis 8-17
8.3.1 Determination of Particle Size Distribution 8-17
8.3.2 Determination of Total Organic Carbon Content 8-19
8.3.3 Determination of Total Metal Concentrations 8-20

8.4 Sediment Monitoring as a tool for Investigation of Cause 8-21
8.4.1 Sediment Mass Transport 8-21
8.4.2 Sediment Depositional Rate 8-22
8.4.3 Sediment Coring 8-23
  8.4.3.1 Recommended Procedures for Core Sampling 8-23
  8.4.3.2 Transport and Storage of Core Samples 8-27
  8.4.3.3 Manipulation of Collected Samples 8-27
8.4.4 Sediment Chemistry 8-28
  8.4.4.1 Partial Metal Concentrations in Sediment (Partial Extraction), Sequentially Extracted Metals, SEM/AVS Ratios 8-28
  8.4.4.2 C:N Ratio for Marine Sediment 8-29
8.4.5 Whole Sediment Toxicity Testing 8-29
  8.4.5.1 Whole Sediment Toxicity Test: Survival and Growth of *Hyalella azteca*8-30
  8.4.5.2 Whole Sediment Toxicity Test: Survival and growth of the freshwater midges Chironomus tentans or Chironomus riparius 8-31
  8.4.5.3 Whole Sediment Toxicity Test: Survival and reproduction of the oligochaete *Tubifex tubifex* 8-31
  8.4.5.4 Whole Sediment Toxicity Test: Survival and growth of the mayfly *Hexagenia limbata* 8-31

8.4.5.5 Using a battery of whole sediment toxicity tests: *Hyalella azteca, Chironomus tentans* or *Chironomus riparius, Tubifex tubifex* and *Hexagenia limbata* 8-32
8.4.5.6 Metals in Overlying Water in Whole Sediment Toxicity Tests........ 8-32
8.4.5.7 Bioaccumulation: Measurement of Metals in Test Organisms from Sediment Toxicity Testing........ 8-33
8.4.6 The Sediment Quality Triad 8-34
8.4.7 Sediment Pore Water Monitoring........ 8-36
8.4.7.1 Collection of Pore Water 8-37
8.4.7.2 Sediment Pore Water Toxicity Testing........ 8-40
8.4.8 Toxicity Identification Evaluation and Toxicity Reduction Evaluation (TIE/TRE)........ 8-41

8.5 References ***8-43***

8.6 Appendix 8.1: Details on In-situ Collection of Pore Water ***8-54***

8.7 Appendix 8.2: Details on Recovery of Pore Water from Bulk or Core Sediment Samples........ ***8-56***

List of tables

Table 8-1: Advantages and disadvantages of the different grab devices most commonly used to collect sediments (ASTM 1992a; Mudroch and MacKnight 1991).............. 8-4
Table 8-2: Minimum volume or weight of sediment required for a specific end use.... 8-10
Table 8-3: Type of container and conditions recommended for storing sediment samples8-15
Table 8-4: Type of container and conditions recommended for storing sediment samples collected for chemical analysis 8-17
Table 8-5: Advantages and disadvantages of the different sediment corers most commonly used to collect sediments (ASTM 1992a; Mudroch and MacKnight 1991)8-25
Table 8-6: A summary of the various methods for the collection of pore water, *in-situ*8-39

List of figures

Figure 8-1: Recommended sediment samplers for different types of freshwater environments (after Environment Canada 1994)........ 8-5
Figure 8-2: Recommended sediment samplers for different types of estuarine environments (after Environment Canada 1994)........ 8-6
Figure 8-3: Recommended sediment samplers for different types of marine environments (after Environment Canada 1994)........ 8-7

8. Sediment Monitoring

Overview

As part of each benthic invertebrate community survey completed for biological monitoring studies, mines are required to collected sediment samples for analysis for particle size distribution and total organic carbon. Sediment samples will be collected at the same sampling stations and at the same time as benthic invertebrate samples. Each study design for benthic invertebrate community surveys will identify the sediment sample collection and laboratory analysis methods to be used (field and laboratory methodologies selected). The results of these analyses will be included in the interpretative report. The results of analyses of particle size distribution and total organic carbon will be used to determine if there are habitat differences between the exposure and reference areas, to aid in the interpretation of the results of benthic invertebrate community surveys. The overall purpose of sediment monitoring is to answer the question: "Are there habitat differences that may contribute to effects in the benthic invertebrate community?".

When a benthic invertebrate community survey is required as part of Focussed Monitoring, it is recommended that each sediment sample collected also undergo chemical analysis. The study design for that study should identify the parameters for which sediment samples will be analysed, and the laboratory methods to be used, and the results of analyses should be reported in the interpretative report.

When Investigation of Cause is required to identify the causes of the effect on the benthic invertebrate community, detailed studies of sediment may be appropriate as a tool to help determine the cause of effect. This chapter contains extensive technical guidance on the conduct of detailed sediment studies recommended as part of Investigation of Cause.

8.1 Collection of Sediment Samples

This section provides guidance on the collection, handling, storage and transportation of sediment samples. This guidance applies to sediment samples collected in all phases of the EEM program.

8.1.1 Field Measurements and Observations

Field measurements and observations are critical to any sediment collection study. It is recommended that the following information (Mudroch and MacKnight 1991) be recorded at the time each sediment sample is collected from a sampling station:

- sample number, replicate number, station number, site identification (e.g., name);
- time and date of the collection of the sample;
- ambient weather conditions, including wind speed and direction, wave action, current, tide, vessel traffic, temperature of both the air and water, thickness of ice if present;
- station location (e.g., positioning information) and location of any replicate samples;
- type of platform/vessel used for sampling (e.g., size, power, type of engine);
- type of sediment collection device and any modifications made during sampling;
- the water depth at each sampling station and the sediment sampling depth;
- name of personnel collecting the samples;
- details pertaining to unusual events which might have occurred during the operation of the sampler (e.g., possible sample contamination, equipment failure, unusual appearance of sediment integrity, control of vertical descent of the sampler, etc.);
- description of the sediment including texture and consistency, colour, odour, presence of biota, estimate of quantity of recovered sediment by a grab sampler, or length and appearance of recovered cores (photographs provide a good permanent record of a retrieved sample); and
- deviations from SOPs
- unusual or unpredicted events

8.1.2 Criteria for Selection of a Sample Collection Device

There are numerous methods and procedures reported in the literature which describe how to collect various types of sediment samples in different types of environments (for reviews see Baudo et al. 1990; Mudroch and MacKnight 1991; ASTM 1992a; Burton 1992). Environment Canada (1994), Baudo et al. (1990) and Håkanson and Jansson (1983) suggest several factors that should be considered for the selection of sediment samplers and sampling location. The ideal sediment sampler should for the most part:

- permit free water passage during descent, to avoid a pressure wave;
- have a sharp-edged cutting surface, a small edge angle, smooth inside surface, and small wall thickness to minimise disturbance;
- close tightly for the ascent;
- allow subsampling;
- have the capability of adjusting weight for penetration of different substrates;
- be able to retrieve a volume of sediment large enough to meet the analytical test requirements;

- effectively and consistently retrieve sediments from various water depths;
- effectively and consistently retrieve sediments from the desired sampling depth;
- not contaminate or influence the nature of the sediment;
- require a minimum of supportive equipment;
- be easy and safe to operate and not require extensive training of personnel; and
- be easily transported to and assembled at the sampling site.

Table 8.1 and Figures 8.1, 8.2 and 8.3 present information to help determine the most appropriate sampling device for a freshwater, marine or estuarine environment. Most sediment samplers are designed to consistently isolate and retrieve a volume of sediment to a required depth below the sediment surface with minimum disruption to the integrity of the sample and no contamination of

Table 8-1: Advantages and disadvantages of the different grab devices most commonly used to collect sediments (ASTM 1992a; Mudroch and MacKnight 1991)

Grab Sampler /Dimensions	Use	Sediment Depth Sampled (cm)	Volume of Sediment Sample (cm^3)	Advantages	Disadvantages
Orange-Peel Grab, Smith-McIntyre Grab	Deep lakes, rivers and estuaries.	0-30	10,000-20,000	Designed for sampling hard substrates	Loss of fine-grained sediment; heavy relative to other grabs; may require winch; possible metal contamination
Birge-Ekman – small	Lakes and marine areas. Soft sediments, silt and sand	0-10	$\leq$ 3400	Designed for fine-grained soft sediments and mixtures of silt and sand; light weight therefore easy to operate manually	Restricted to low current conditions
Birge-Ekman – large	Lakes and marine areas. Soft sediments, silt and sand	0-30	$\leq$ 13,300	Designed for fine-grained soft sediments and mixtures of silt and sand; large sample obtained permitting subsampling	Restricted to low current conditions; penetration depth exceeded by weight of sampler in very soft sediment
PONAR Grab - standard	Deep lakes, rivers and estuaries. Useful on sand, silt or clay	0-10	7,250	Most universal grab sampler; adequate on most substrates; large sample obtained intact, permitting subsampling; good for coarse and firm bottom sediments	Shock wave from descent may disturb fine-grained sediment; possible incomplete closure of jaws results in sample loss; possible contamination form metal frame construction
PONAR Grab – mini	Deep lakes, rivers and estuaries. Useful on sand, silt or clay	0-10	1,000	Adequate for most substrates that are not compacted	Smaller volume does not minimise disturbance to sample
Van Veen	Deep lakes, rivers and estuaries; useful on sand, silt or clay; effective in marine environments in deep water and strong currents	0-30	18,000-75,000 18-75 L	Adequate on most substrates; large sample obtained intact, permitting subsampling; available in stainless steel	Shock wave from descent may disturb fine-grained sediment; possible incomplete closure of jaws results in sample loss; premature closing in rough waters; possible contamination from metal frame construction
Petersen Grab Sampler	Deep lakes, rivers and estuaries. Useful on most substrates	0-30	9,450	Large sample; can penetrate most substrates	Heavy, likely requires winch; no cover lid to permit subsampling; all other disadvantages of Ekman and Ponar
Shipek Grab Sampler - standard	Used primarily in marine waters and large inland lakes and reservoirs; not useful for compacted sandy clay or till substrates	0-10	3,000	Sample bucket may be opened to permit subsampling; retains fine-grained sediments effectively	Possible contamination from metal construction; heavy, may required winch
Shipek Grab Sampler -mini	Lakes, useful for most substrates that are soft	0-3	500	Easily operated by hand from most platforms	Requires vertical penetration; small volume; washout of fine-grained sediment; premature closing

Figure 8-1: Recommended sediment samplers for different types of freshwater environments (after Environment Canada 1994)

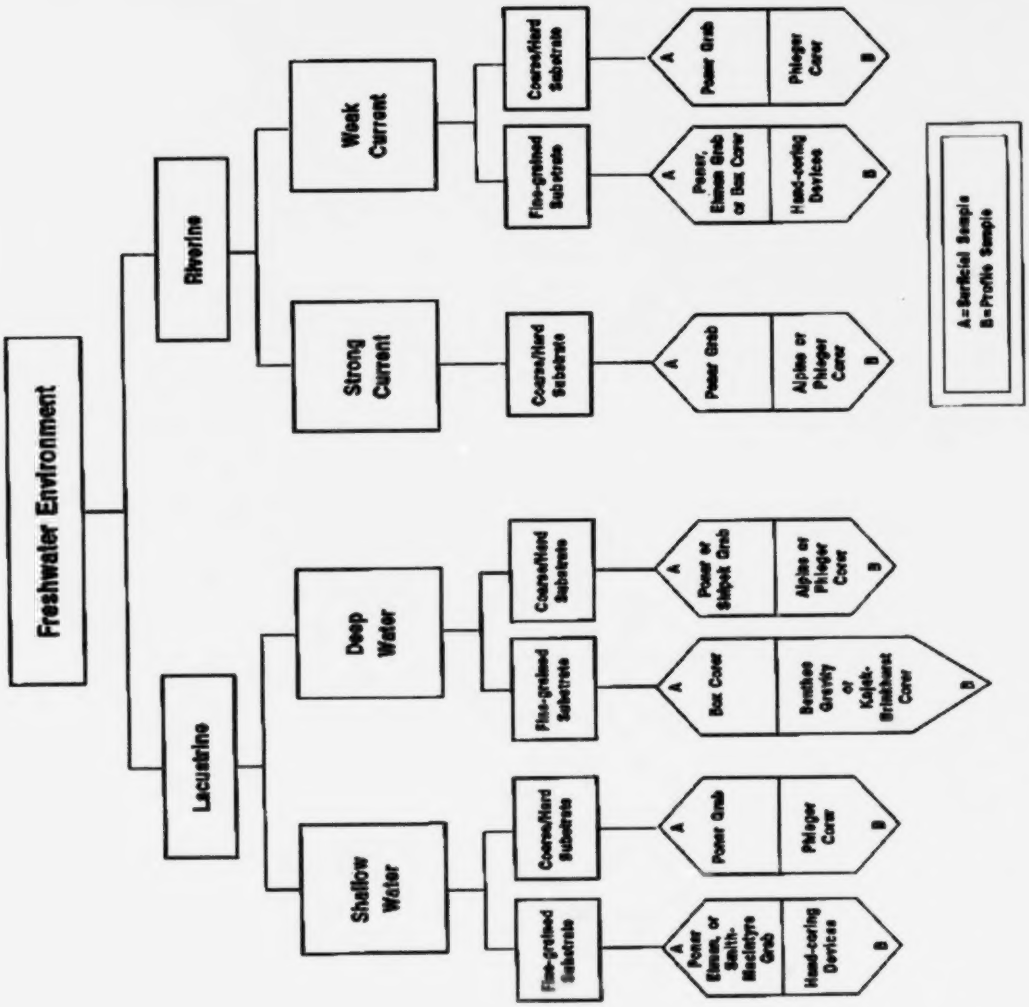

Freshwater Environment
Riverine
Week Current
Coarse/Hard Substrate
A Ponar Grab
Phleger Corer B
Fine-grained Substrate
A Ponar, Ekman Grab or Box Corer
Hand-coring Devices B
Strong Current
Coarse/Hard Substrate
A Ponar Grab
Alpine or Phleger Corer B
Lacustrine
Deep Water
Coarse/Hard Substrate
A Ponar or Shipek Grab
Alpine or Phleger Corer B
Fine-grained Substrate
A Box Corer
Benthos Gravity or Kajak-Brinkhurst Corer B
Shallow Water
Coarse/Hard Substrate
A Ponar Grab
Phleger Corer B
Fine-grained Substrate
A Ponar Ekman, or Smith-McIntyre Grab
Hand-coring Devices B
A=Surficial Sample
B=Profile Sample

Figure 8-2: Recommended sediment samplers for different types of estuarine environments (after Environment Canada 1994)

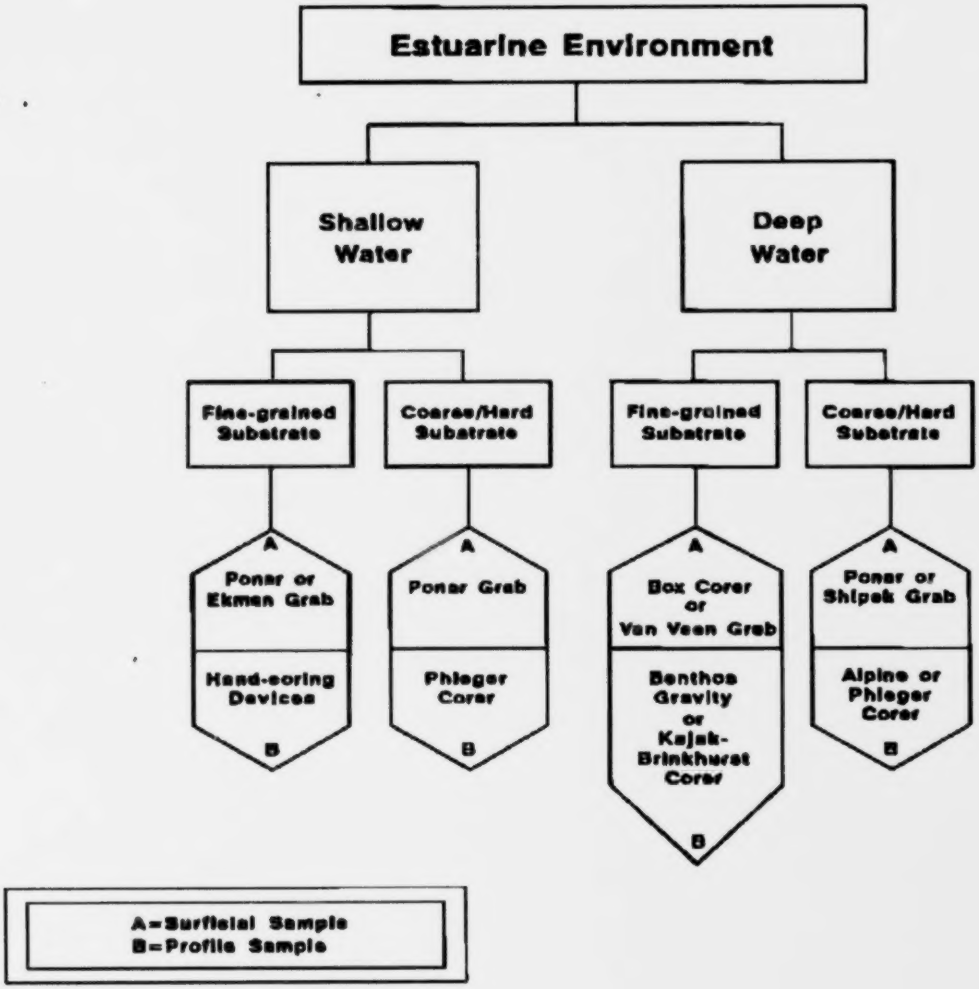
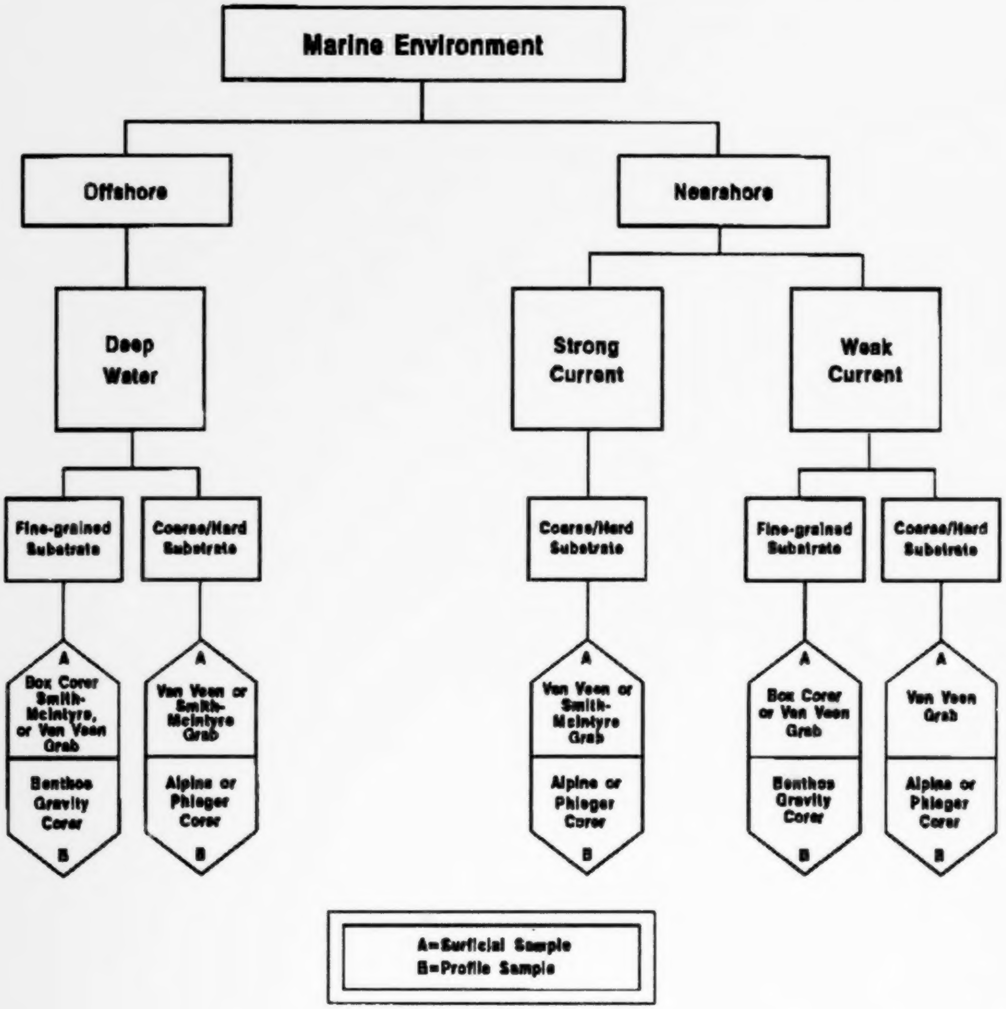

Estuarine Environment
Shallow Water
Deep Water
Fine-grained Substrate
Coarse/Hard Substrate
Fine-grained Substrate
Coarse/Hard Substrate
A
Ponar or Ekman Grab
Hand-coring Devices
B
A
Ponar Grab
Phleger Corer
B
A
Box Corer or Van Veen Grab
Benthos Gravity or Kajak-Brinkhurst Corer
B
A
Ponar or Shipek Grab
Alpine or Phleger Corer
B
A=Surficial Sample
B=Profile Sample

Figure 8-3: Recommended sediment samplers for different types of marine environments (after Environment Canada 1994)


Marine Environment
Offshore
Nearshore
Deep Water
Strong Current
Weak Current
Fine-grained Substrate
Coarse/Hard Substrate
Coarse/Hard Substrate
Fine-grained Substrate
Coarse/Hard Substrate
A
Box Corer Smith-McIntyre, or Van Veen Grab
Benthos Gravity Corer
B
A
Van Veen or Smith-McIntyre Grab
Alpine or Phleger Corer
B
A
Van Veen or Smith-McIntyre Grab
Alpine or Phleger Corer
B
A
Box Corer or Van Veen Grab
Benthos Gravity Corer
B
A
Van Veen Grab
Alpine or Phleger Corer
B
A=Surficial Sample
B=Profile Sample

the sample. Maintaining the integrity of the collected sediments is of primary concern in most studies, since disrupting the structure of the sediment may change the physicochemical and biological characteristics, which in turn could influence the partitioning, complexation, speciation and bioavailability of the toxicants. Sometimes it is also important to maintain the profile if sectioning is required at different depths. These issues become even more important during Investigation of Cause Monitoring studies when sediment may be collected for toxicity tests or more complex analytical methods (see Section 8.4). In general, it is recognised that it is difficult to collect a sediment sample with most sampling devices without some degree of disruption.

There are three main types of sediment samplers; grab, core and dredge samplers. For Initial, Periodic and Focussed Monitoring grab samplers are recommended. Corers should be reserved for Investigation of Cause , and are further discussed in section 8.4. Grab samplers are used to collect surficial sediments for the determination and assessment of the horizontal distribution of sediment characteristics. The advantages and disadvantages of the various grab samplers are summarised in Table 8.1 and discussed briefly below. Details on this topic can be found in de Groot and Zschuppe (1981), Baudo et al. (1990), ASTM (1992a), Burton (1992) and Sly and Christie (1992).

Grab samplers have the advantage of being easy to handle and operate, readily available, moderately priced, versatile in terms of substrate type, and they can collect either small volumes (0-10 cm deep; e.g., Birge-Ekman, Ponar, mini-Ponar, or mini-Shipek) or large volumes (0-30 cm deep; e.g., Van Veen, Smith-McIntyre, Petersen). Careful use of these sampling devices avoids most of the problems associated with unpredictable penetration of sediment, loss of sediment from tilting or washout upon ascent, mixing of sediment layers upon impact, loss of fine-grained surface sediments from the bow wave during ascent, and susceptibility to the influences of waves and currents (Plumb 1979; Golterman et al. 1983; Blomqvist 1990; Baudo et al. 1990; ASTM 1992a).

Dredges are used primarily for the collection of benthos, since they are usually equipped with net sides designed to filter out fine-grained sediments and retain coarse sediments and fauna. It is virtually impossible to accurately measure the surface area covered by the dredge sampler, or judge the depth to which the sediment sample has been collected. In addition, sediment integrity is disrupted, pore water excluded, and fine-grained sediments lost during ascent using dredge samplers. For these reasons, only grab (Initial, Periodic and Focussed Monitoring) and core samplers (Investigation of Cause) are being recommended for collection of sediments.

8.1.3 Collection Device Penetration Depth

The desired depth of sediment penetration is a decision that depends upon the type of sampling device, the nature of the sediment, and the volume of sediment required. The actual depth of penetration depends primarily on the type of sampling device and the nature of the sediment. For most monitoring studies where historical contamination is not a priority the upper 0-5 cm of sediment is the horizon of interest for chemical characterization and toxicity evaluation. Generally, the most recently introduced contaminants of concern and most infaunal organisms are found in the upper 2 cm, and epifaunal organisms have access to this horizon (Burton 1992). Therefore, a preferred penetration depth of 10-15 cm and a minimum penetration depth of 6-8 cm are recommended to ensure minimum disturbance of the upper layer during sampling.

8.1.4 Sample Size (volume of sediment)

The recommended minimum volume or weight of sediment required for each end use is summarised in Table 8.2. Accordingly, the sample size should be determined on a case-by-case basis. Before commencing a sampling program, the type and number of analyses and tests should be determined, and the required volume or weight of sediment per sample calculated. Each physico-chemical test requires a specific amount of sediment. After the sample size is determined, it is important to compare the sample size required with the capacity of the sampler to deliver the desired amount of sediment (Table 8.1), and reassess the number of replicate samples per station. The volume or weight requirements might dictate further sample handling such as subsampling, compositing, or sample splitting.

8.1.5 Replicate Samples

When conducting benthic invertebrate community survey, at least one sediment sample will be collected at each benthic sampling station in reference and exposure areas. Most monitoring studies have traditionally collected only one sample from each sampling station. However, a single sediment sample from a sampling station will impart little information on the variability in the sediment. The objective of collecting separate replicate samples at each sampling station is to allow for quantitative statistical comparison within and among different stations (Holland et al. 1993). Separate subsamples from the same grab might be used to measure the variation within a sample but not necessarily within the sampling station. The collection of separate samples within a sampling station can impart valuable information on the heterogeneity of the sediments.

The number of replicates required per sampling station is a function of the need for sensitivity or statistical power. Typically, the smallest deviation from the null hypotheses that is considered scientifically or environmentally important to detect should be decided *a priori*, together with the power of the test that is desired for the specific alternative (Green 1989).

8.1.6 Grab Sampler Operation

When collecting bottom sediments with grab samplers, the speed of descent of the sampling device should be controlled and the sampler should not be permitted to "free fall". To minimise twisting during the descent, a ball bearing swivel should be used to attach the sampler to the cable. The sampler should contact the substrate or be positioned just above it and only its weight or piston mechanism should be used to force it into the sediment. The winching system should be in place to control both the ascent and descent of the sampling device, especially in deep water. After the sample is contained, the sampling device should be lifted slowly off the bottom, then steadily raised to the surface at about 30 cm/s. When the sampler is brought to the surface, the outside of the sampler should be carefully rinsed with water from the sampling station to remove material that could potentially contaminate the sample during transfer, and inspected to ensure that the sampler has closed properly. The standard operating procedures specific for each grab sampler should be followed in order to ensure proper operation of the sampler.

Table 8-2: Minimum volume or weight of sediment required for a specific end use

End Use	**Volume (mL)**	**Weight[1]**	
		(g dry weight)	**(g wet weight)**
Physical/Chemical Analyses:			
Inorganic Contaminants	90	10	100
Other Chemical Constituents (e.g., TOC, moisture content)	300	60	330
Particle Size	230	50	250
Biological Testing			
Toxicity Tests[2]	1000 – 3000	200 – 600	1100 - 3300
Bioaccumulation Tests[3]	3000	600	3300
Pore water Extraction[4]	2000	NA	3200
Preparation of Elutriate	1000	200	1100

[1] Based on a specific gravity of 2.0, a moderate organic matter content, and a water content of 90%

[2] Based on the average requirement for three tests

[3] Based on an average of 3 L of sediment per sample; an additional 3 L is required for each replicate

[4] Based on a specific gravity of 2.0, a moderate organic matter content, and an extractable water content of 40%

NA Not applicable

Regardless of the type of samplers used, standard operating procedures for each device should be immediately accessible, and all personnel involved with the collection of samples should be familiar with these procedures. The sampling vessel or platform should be stationary, and sufficiently stable to permit inspection and handling of the retrieved sample. Field notes should accompany each sample that is collected. The sampling device should be cleaned thoroughly between sampling stations and between within-station samples by dipping the sampler into and out of the water at a rapid speed to wash the sediment off. Alternatively, a hose can be used to wash the sediment off of the sampler with water from the sampling station. The sampler should be rinsed with water from the next sampling station before collecting a sample.

Summary of recommended operations and procedures for grab devices:

- the grab samplers recommended for the collection of depositional sediments in freshwater, estuarine, and marine environments are presented in Figures 8.1, 8.2 and 8.3, respectively;
- a minimum penetration depth of 6-8 cm is recommended for surficial sediment sampling; however, a depth of 10-15 cm is preferred;
- appropriate winching systems are required to control the rate of ascent and descent of the samplers in deeper water;
- the sampling vessel or platform should be stationary, and as stable as possible;
- field notes and/or measurements should accompany each sample;
- if sediments adhere to the outside of the sampler, the external surface of the sampler should be carefully washed/rinsed with clean water upon retrieval before the sample is transferred to a storage container;
- the sampler should be rinsed thoroughly with water at the sampling station between within-station samples, and rinsed with water from the next sampling station before collecting a sample. Equipment used in the handling of sediment should also be washed thoroughly between samples;
- a sample should meet the criteria of acceptability (Section 8.1.7) before it is considered adequate.

8.1.7 Criteria of Acceptability of Samples

All samples should be visually inspected to ensure that:

- the desired depth of penetration has been achieved;
- there is no evidence of incomplete closure of the grab sampler, or that the grab sampler was inserted on an angle or tilted upon retrieval (i.e., loss of sediment);

If the collected sample fails any of the criteria listed above, then the sample should be rejected and another sample collected at the site. The location of consecutive attempts should be as close to the original attempt as possible and, where the direction of the current is known, consecutive attempts should be located in the opposite direction of the current or "upstream". Rejected sediment samples should be discarded in a manner that will not affect subsequent samples at that station or other possible sampling locations.

8.1.8 Sediment Parameters that May be Measured in the Field

It is recommended that the following sediment parameters be measured in the field, particularly during Focussed Monitoring and Investigation of Cause:

- temperature and pH of the sediment at the sediment-water interface;
- a measure of the redox potential of the sediments to determine if the sediments are oxic or anoxic, or to determine the depth of the interface between these conditions in the sediments. Dissolved oxygen is recommended for freshwater sediment and Eh is recommended for marine sediments

These measurements could be useful for the interpretation of the analytical results.

8.2 Sample Handling and Analysis

8.2.1 Procedures for Handling of Sediment Samples

Any time that sediment samples are handled, it is recommended that the following procedures be observed:

- sediment might contain a mixture of hazardous substances, so it is prudent to avoid skin contact with sediments by wearing protective clothing and equipment (e.g., gloves, boots, lab coats or aprons, safety glasses, and respirator) during sampling, sample handling, and the preparing of test substances.

- handling of samples should be performed in a well-ventilated area (e.g., outside, in a fume hood, or in an enclosed glove box) to minimise the inhalation of sediment gases.

- work surfaces should be covered with Teflon® sheets, high-density polyethylene trays, or other impervious or disposable, similarly inert material.

- a spill control protocol should be in place in the laboratory or sampling vessel, and participants in the project should be familiar with all SOPs and recommendations.

8.2.2 Subsampling of Sediment Grab Samples

If sediment grab samples are to be subsampled, then access to the surface of the sample without a loss of water or fine-grained sediment is a prerequisite for selection of the sampler.

The non-turbid overlying water, if present, should be gently siphoned off before the sediment is subsampled, using a flat, clean scoop (e.g. Teflon® or a similarly inert, non-contaminating, non-reactive material) or a suitable hand-coring device. The sediment should be collected to a depth of 5 cm. Subsampling the top 2 cm is a common practice when the most recent sediment deposits are of concern, but the actual depth chosen depends on the objectives of the study. Ideally, each subsample should be placed into a clean, separate, pre-labelled container. The labelled sample container should be sealed and the air excluded.

In the event that the collection device does not allow access to the surface, the following procedures should be followed. Upon retrieval of the sample, the contents should be carefully deposited into a clean, inert container that is the same shape as the sampler. The sampler is placed into the container and the jaws opened slowly to allow the sample to be deposited into the container with as little disturbance as possible. Once the sample is in the container, subsamples can be collected from the sample with a hand corer or scoop. The edges of the sample where the sediments may be disturbed during removal from the sampler should be excluded during subsampling.

8.2.3 Compositing Sediment Grab Samples

If the objective of the study dictates compositing subsamples from separate grabs within a sampling station, then the subsamples may be placed into one clean sample container and, when full, sealed without trapped air. Compositing of sediment samples or subsamples may also be performed in the laboratory.

8.2.4 Sample Containers

Sediment samples should be transferred from the sampler to sample containers that may or may not serve as the storage container as well. The containers can be stored temporarily in the field before being transported to and stored at the laboratory, or they can be transported immediately to the laboratory for storage. The recommended sample containers for analysis of particle size and total organic carbon are identified in Table 8.3. Whole-sediment samples may be transferred directly from a sampler into a clean, large volume (e.g. >1 L) container. If smaller volumes of sediment are collected or subsampled, then containers with wide mouths and Teflon®-lined lids are recommended for volumes ranging from 250 to 1000 mL.

If samples are to be stored at 4°C, then sample containers should be filled to the rim and air excluded during capping. If samples are to be frozen for storage then glass containers should not be filled completely. A space of approximately 2.5 cm should be left to accommodate expansion of the sample when frozen. The headspace in the container should be purged with nitrogen before capping tightly. Clear glass containers may be wrapped with an opaque material (e.g. clean aluminium foil) to eliminate light and reduce accidental breakage.

Each sample container should be properly labelled and stabilised in an upright position in the transport container. Labelling of each sample container should include, at a minimum, the site, station location or identification, the sample type, the method of collection, the name of the collector, and the date and time of collection.

8.2.5 Transportation and Storage of Sediment Samples

The recommended procedures and conditions for the transportation and storage of sediment samples are as follows:

- the transport container should be refrigerated to 4 ± 2°C or contain ice or frozen gel packs that will keep the field samples below 7°C during transport to the laboratory;
- if field-collected samples are warm (e.g. > 6°C), they should be cooled to between 1 and 6°C with ice prior to placement in the transport container;
- samples should not freeze during transport;
- ideally, a max/min thermometer or a continuous temperature recorder should be placed inside the transport container and the container sealed. Deviations in temperature should be reported;
- light should be excluded from the transport container; and,
- all field-collected samples that require further processing before storage should be transported to the laboratory within 72 h, preferably within 24 h, of collection.

Where these conditions cannot be met due to operational constraints, then the storage method and conditions adopted should strive to compromise the integrity of the sample as little as possible (Mudroch and MacKnight 1991).

Table 8-3: Type of container and conditions recommended for storing sediment samples

End Use	Container Type	Wet Weight or Volume of Sample	Storage Conditions	
			Temperature	Holding Time
Particle Size Distribution	Teflon®Glass High-density polyethylene containers or bags	250 g	4 - 40°C Do not freeze	≤6 mo
Total Organic Carbon	Teflon®Glass with Teflon®or Polyethylene-lined cap	100 g	4 ± 2 °C	≤ 48 h

8.2.6 Laboratory Test Sample Preparation

Sediment samples should be prepared in a well-ventilated area (e.g., fumehood) and the appropriate health and safety precautions should be followed. For Initial, Periodic and Focussed Monitoring the preparation of samples under anoxic conditions is not a concern. However, for Investigation of Cause techniques such as toxicity testing, preparation of anoxic test sediments should be performed in a glove box in the presence of a controlled flow of an inert gas, if it is desirable to maintain these anoxic conditions. Below are some details on sediment preparation techniques used to allocate sediment to test containers:

Homogenising: Mixing by hand or mechanical means may be used to achieve homogeneity of colour, texture and moisture; however, the efficacy of the method should be demonstrated, *a priori*, and the mixing time standardised to ensure consistency and minimise alterations in the size distribution of sediment particles, respectively.

Mixing of sediments should take place in the sample/storage container.

Partitioning: Coning or caking and quartering are the recommended techniques for partitioning the sediment for distribution among test containers. If a sediment splitter is used, its efficacy should be demonstrated and documented and it should be made of an appropriately inert material.

Drying: Oven drying sediment subsamples (1-5 g of wet sediment) at low temperatures (40-60°C) until a constant weight is reached or freeze drying sediment subsamples are the recommended methods for drying sediment.

Crushing/Grinding: Commercially-available ball and pebble mills are recommended for fine-grinding small volumes of sediment (Mudroch and MacKnight 1991); however, it should be noted that grinding could change the chemistry of the material. Crushing can usually be achieved with a mortar and pestle.

Dewatering: Centrifugation with subsequent decanting of the supernatant is the recommended method for dewatering sediment samples. The centrifugation speed depends on the sample size and particle size (e.g., sediment weight or volume).

8.2.7 Prevention of Sediment Sample Contamination

When sediment samples are to be collected for chemical analysis the procedures for the collection, handling, transportation and storage of samples are much the same as those outlined above. However, in such cases it is important that appropriate measures are taken to ensure that sediment samples are not contaminated.

When sediment samples are to be collected for chemical analysis sample collection devices should not be made of copper, zinc, brass or galvanised material, since collection devices made of unprotected metallic material can potentially affect the concentrations of metals in sediment samples. If this is not possible, then, when subsampling, the sediment that is in direct contact with the sides of the sampler should be excluded. The subsampled sediment should be transferred to clean containers comprised of inert material which will neither contaminate nor influence the characteristics of the sediment sample. The container should be tightly sealed and air should be excluded.

Types of sample containers and conditions recommended for storing sediment samples collected for chemical analysis are described in Table 8.4. All sample containers should be pre-treated prior to receiving a field sample (Environment Canada 1983, 1989). New glass and most plastics should be pre-treated to remove residues, and/or leachable compounds, and to minimise potential sites of adsorption. Pre-treatment includes the following sequence of activities (adapted from Environment Canada 1989):

- scrub with phosphate-free detergent and hot water;
- rinse with high-pressure hot water;
- subject to a 72-h acid bath with 8 M HNO_3 (50 mL of HNO_3 per litre of water);
- rinse four (4) times with hot water;
- rinse three (3) times with DDW (Double Distilled Water); and,
- wash bottle caps (Teflon® or Teflon®-lined) with detergent and hot water, and rinse with DDW.

The acid bath will leach trace metals (e.g., Cu, Fe, Mo, Ni, Zn) from plastics. The triple rinse with distilled water is necessary because the acid treatment can activate adsorption sites on polymers which are then capable of binding trace metals in the field sample.

Table 8-4: Type of container and conditions recommended for storing sediment samples collected for chemical analysis

End Use	Container Type	Wet Weight or Volume of Sample	Storage Conditions	
			Temperature	Holding Time
Major ions and elements: e.g., Al, C, Ca, Cl, Cr, Fe, Fl, H, K, Mn, Mg, Na, P, S, Si, Ti, (oxides and total)	Teflon® High-density polyethylene containers or bags	250 g	4 ±2°C	≤ 2 wk
Nutrients: e.g., NH_4-N, NO_2-N, NO_3-N, TKN, TC, DOC	Teflon® Glass with Teflon® or Polyethylene-lined cap	100 g	4 ± 2 °C	≤ 48 h
Trace elements: Ag, Ba, Be, Cd, Co, Cr, Cu, Hg, Li, Mn, Mo, Ni, Pb, Sb, Sr, Va, Zn	Teflon® High-density polyethylene containers or bags	250-500 g	4 ± 2°C or - 20°C	≤ 2 wk ≤6 mo

8.3 Sediment Sample Analysis

8.3.1 Determination of Particle Size Distribution

The determination of sediment particle size distribution is required each time that a benthic invertebrate community survey is conducted. Particle size is to be determined for a minimum of one sample from each benthic sampling station.

Particle size determination is important for several reasons. Knowledge of the particle size distribution of sediment is very important in the interpretation of the results of chemical or biological analyses. Most importantly, from the point of view of using this data to aid in the interpretation of the results of benthic invertebrate community surveys, particle size has a significant impact on the structure of benthic invertebrate communities. It may also provide insight into the origin of sedimentary materials and about the dynamic conditions of sediment transport and deposition. From particle size analysis, specific surface, expressed as m^2/g, can be determined, and with this, the adsorptive capacity of metals and organic substances can be assessed.

Many different classifications of particle sizes exist; however, the following breakdown based on the Wentworth Classification (1922) is recommended for the interpretation of

EEM data.

Classification	Particle Size
Gravel	2.0 mm – 16.0 mm
Coarse Sand	2.0 mm – 0.2 mm
Fine Sand	0.2 mm – 0.062 mm
Silt	0.062 mm – 0.0039 mm
Clay	< 0.0039 mm

Samples for particle size analysis can be collected in plastic, glass or metal containers, however, plastic is preferable. Samples should be stored in the dark at 4ºC and analysed within 14 days of collection.

Procedures for methods of sediment particle size analysis can be found in "Standard Method for Particle-Size Analysis of Soil", ASTM, D422-63. Particle size analysis or grain size analysis is generally performed in two parts, sieve analysis and hydrometer analysis. The sieve analysis classifies particles greater than 0.06 to 0.075 mm in size (actual minimum size depends on the sieve set used). This is done by wet sieving the sample through a set of at least four sieves, ranging in size from 0.06 mm to 16 mm. The material retained on the sieves is dried and weighed. Particles passing through the 0.06 mm sieve are collected and transferred to a 2-L container, together with the wash water. A hydrometer is used to determine the quantity of particles in this fraction from 0.06 mm down to 0.0014 mm. The data from these two tests is then tabulated and calculated to produce a particle size distribution curve. This curve graphically defines the percentage of material in the different fractions based on the total sample weight.

It is also possible to determine particle size distribution using laser diffraction, and this method is increasingly available. This method is more efficient, and provides higher resolution results than the above methods. A laser diffraction instrument uses light from a low power helium-neon laser (the analyser beam). Particles from sediment samples enter the beam via a dispersion tank that pumps the material, carried in water, through a sample cell. The light scattered by the particles is incident onto the receiver lens, which focuses the scattered light onto a diode composed of numerous concentric rings. Through a process of constrained least squares fitting of theoretical scattering predictions to the observed data, the computer calculates a volume size distribution that would give rise to the observed scattering characteristics. No *a priori* information about the form of the size distribution is assumed, allowing for the characterization of multi-modal distributions.

The efficiency of laser diffraction is also a major benefit. A typical measurement takes only a few seconds, and the data are saved digitally, and are instantly available for plotting and other calculations. Often, the entire distribution can be accounted for in a single measurement. Depending on the instrument used, a laser particle size analyser can measure all sizes ranging from 0.05µm - 2000µm. For samples with a size range greater than 2000µm, sieve data can be merged with the laser results. Finally, the results using laser diffraction are very high resolution, and are easily reproducible – a major shortcoming of the hydrometer and sieve methods.

8.3.2 Determination of Total Organic Carbon Content

Like particle size distribution, the determination of sediment total organic carbon is required each time that a benthic invertebrate community survey is conducted. Total organic carbon is to be determined for a minimum of one sample from each benthic sampling station.

Carbon is present in sediment in several organic forms such as humic matter, chemical, plant and animal matter as well as inorganic carbonate forms. Organic carbon in sediment and the water column causes a decrease in dissolved oxygen by using up available oxygen, hence creating a more eutrophic environment. Also, at certain pH levels, humic substances form complexes with metals, increasing metal solubility in the water column. Two methods are commonly used to analyse total organic carbon (TOC) in sediment. The elemental analyser method, valid for 0.5 mg to 25 mg samples, is based on the use of thermal conductivity. The oxidising furnace method requires 0.25 g to 0.5 g samples and is based on the use of infrared spectrophotometry.

Elemental analyser: Inorganic carbon is first eliminated by treatment with hydrochloric acid. TOC is then oxidised to carbon dioxide in the presence of a catalyst. The gas produced is separated by chromatography and quantified with a thermal conductivity detector.

Oxidising furnace: Inorganic carbon is first eliminated by treatment with hydrochloric acid. TOC is then oxidised in the oxidising furnace in the presence of manganese dioxide. The carbon dioxide formed from the organic carbon is measured directly by infrared absorption at the characteristic wavelength for carbon dioxide.

Procedures for these methods of analysing TOC in sediment are described in:

Test Methods for Evaluating Solid Waste-Physical and Chemical Methods, USA Environmental Protection Agency, 1986, Method 9060.

APHA. 1995. Standard Methods for Examination of Water and Wastewater. 19th edition, published by the American Public Health Association (APHA).

8.3.3 Determination of Total Metal Concentrations

The determination of total metal concentrations in sediments is not required as part of the EEM program. However, mines are encouraged to determine total metal concentrations in sediments when benthic invertebrate community surveys are conducted. Information regarding metals in sediment can be important to the interpretation of the results of benthic invertebrate community surveys, and to the design of subsequent surveys.

Sediments are an integral component of aquatic ecosystems, and hence, a frequent aspect of many environmental monitoring programs. They originate from the differential settling of both suspended terrigenous particles that have been introduced into aquatic ecosystems and precipitates that have resulted from chemical and biological processes within aquatic systems. Suspended particles entering the aquatic system may already contain contaminants. Alternately, non-contaminated particles suspended in water may accumulate soluble contaminants present in the waters of aquatic systems. Precipitation processes are also capable of scavenging contaminants. As a result, sediments can be viewed as either a reservoir or a sink for contaminants.

Contamination of sediments from point source inputs such as mining effluents can become bioavailable and enter aquatic food-webs, therefore affecting the quality of the habitat. Measuring sediment quality helps identify which contaminants are entering the exposure area. Sediments provide a better integrator of average long-term environmental condition than single event water chemistry samples.

The selection of parameters for sediment chemistry analysis will be determined on a site-specific basis. Where historical data exists on sediment quality, it should be used in conjunction with effluent characterization and water quality data to help determine parameters to analyse.

Total, or bulk, sediment chemistry provides information on the loading rates of particular elements, and depositional patterns. Techniques used for determination of metals in sediments include atomic absorption spectrophotometry (AAS), X-ray fluorescence (XRF), instrumental neutron activation analysis (INAA), inductively coupled atomic absorption spectrophotometry (ICP-AES) and ICP-mass spectrometry (ICP-MS). Because of the high concentrations of metals in sediment, particularly in mining areas, analytical techniques with higher detection limits (e.g. ICP, ICP-AES) are generally acceptable for sediment chemistry analysis.

Bulk sediment samples are digested using either *aqua regia* or a mixture of perchloric, nitric and hydrochloric acids for extraction of total metals. Metal concentrations can be influenced by sediment particle size and organic carbon content. Smaller particles and organic material have a higher affinity and more binding sites for metals than coarser grained material. Therefore, with all other factors being equal, total metal concentrations tend to be higher in fine organic substrates. To account for this influence, it is recommended that sediment samples be sieved through a 63 µm mesh screen, and that only the less than 63 µm fraction be analysed.

Alternatively, metal concentrations in sediments can be normalised for particle size or organic content when comparing results between areas. Sediment metal data can be normalised to percent fines using the following equation (ESP 1996):

$$\text{Metal}_{NF} = \frac{\text{Metal}}{\text{Fines}}$$

where:

Metal_{NF} = metal concentration normalised to fines
Metal = reported metal concentration in sediment (mg/kg)
Fines = proportion of fines (silt + clay fractions) in sediment

8.4 Sediment Monitoring as a tool for Investigation of Cause

Investigation of Cause occurs when a mine related effect has been identified and previous monitoring studies have failed to provide a satisfactory explanation of the cause of the effect. However, at sites where there are effects in the benthic invertebrate community, there are a number of sediment monitoring techniques which may help in understanding the nature and cause of effects. In some cases, some of these techniques could also provide valuable information during Focussed Monitoring when determining the magnitude and geographic extent of an effect observed in the benthic invertebrate community.

The techniques described below are recommended. Not all these techniques may need to be used, possibly one or few techniques should be applied to investigate the cause of an effect. The following sections describe a range of tools and techniques that could be used to examine sediments during Investigation of Cause.

8.4.1 Sediment Mass Transport

In certain stream systems, an understanding of sediment mass transport characteristics can be important, particularly to understanding the origin and fate of sediments. The morphology of alluvial stream and river channels with mobile beds is believed to be determined by storm events. Thus, the pool–and–riffle sequence characteristics of such river beds, and the corresponding sediment particle size distributions within the channels, may have life-times which are short relative to the EEM program. In these circumstances, appropriate characterization of stream channel morphology, and in particular of its more dynamic elements, may be important information in determining the cause of effects on benthic invertebrate communities. It may be important to know, for example, whether annual deposition and subsequent removal of fine-grained material is a regular feature of the study area, and whether ecologically significant movement of relatively coarse bed material (gravel and cobbles) is taking place on time-scales of a few years or less.

Much of the evidence relating channel morphology to sediment transport and basin hydrology has been gathered in the United States from river basins which yield relatively large amounts of suspended matter. In glaciated regions typical of much of Canada, alluvial sediments in stream channels dominated by material of glacial origin may be more stable in relation to storm events. Further, in the Canadian Shield in particular, the abundance of lakes and wetlands, particularly in first and second order catchments, may help to buffer the effects of storm events. Thus, comparison of results with results of studies in the U.S. may be of limited value.

Direct observations of suspended sediment transport and of bedload movement are difficult because the key events are of short duration and occur at times of stream flows which prevent or render hazardous attempts at direct measurement. However, there are a number of qualitative indicators of the dynamic nature of river channels can be employed. Photographic records of peak run-off events which involve bankfull discharge (e.g. with a return period of the order of a year and greater) can be used to provide qualitative evidence of the importance of suspended sediment transport and deposition. Another method is to identify boulders or cobbles with colour codes and periodically observing the extent of their displacement. It is suggested that the mine operator use a combination of methods to describe the physical setting of such dynamic river systems.

8.4.2 Sediment Depositional Rate

An understanding of the sediment depositional rate can be important in understanding influences on sediment chemistry, and in particular, the relative influence of a mining operation.

Depositional rates may be estimated quantitatively using sediment traps, but representative sampling may be difficult, and it may not be possible, using this method, to account for the impact of storms or other significant hydrological events. Because of these difficulties, direct measurement of sediment deposition rates should be reserved for exceptional circumstances. Long term depositional rates can also be estimated quantitatively using sediment coring.

The relative influence of the mining operation can also be estimated in a more qualitative manner in gravel-bed streams or in lake sediments dominated by soft organic matter. In such cases, the presence of fine mineral-derived sediments would likely indicate some influence from the mine. In other environment types, this method may not be effective, since the deposition of natural mineral-derived fine sediments would mask any influence from the mine.
See Håkanson and Jansson (1983) for guidance on lake sedimentology.

8.4.3 Sediment Coring

Note that the Environment Canada document "Guidance Document on Collection and Preparation of Sediments for Physicochemical Characterization and Biological Testing" (1994) contains extensive guidance on sediment coring. Those considering collecting sediment samples using coring are encouraged to consult that document.

Core samplers are often used to collect sediment profiles for the determination of the vertical distribution of sediment characteristics. Corers are also generally preferred where maintaining the integrity of the sediment profile is essential, because they are considered to be the least disruptive. Core samplers should be used where it is important to maintain an oxygen-free environment below the surficial sediment, since there is less of a chance for oxidation to occur. A range of sediment coring devices are available. Table 8.5 lists a number of coring devices, their uses, and their advantages and disadvantages. Although core samplers have the advantage of collecting minimally disturbed, intact sediment samples from both surficial sediments (upper 15-30 cm) and deep sediments (> 30 cm deep), there are few that function efficiently in substrates with large amounts of sand, gravel, clay, or till. The major limitation of core samplers is that although the core tubes can vary in diameter (3.5-10 cm inside diameter), the volume of the upper surficial horizon (0-5 cm deep) is relatively small, and repetitive, time-consuming sampling may be required to obtain the desired quantity of material. Loss of the highly porous layer at the sediment-water interface due to material displacement during the initial contact, as well as compression during penetration, over penetration, tilting, and loss of sediments during ascent are commonly encountered problems that should be overcome or minimised by ensuring that the proper techniques for collecting the sample are followed.

8.4.3.1 Recommended Procedures for Core Sampling

If the retrieved sediments are core samples to be transported directly to the laboratory intact, the following procedures and precautions should be followed.

- the core liners should be capped or stoppered, taped and closed tightly immediately upon retrieval to prevent loss of sediment and water.
- visual inspection of the intact core should be recorded in field notes, describing the length of the sediment core; the thickness of the various sediment units; the colour, consistency (e.g., packed, loose, consolidated, unconsolidated), texture (e.g., gravel, sand, silt, clay), and presence of shells, organic matter, vegetation, organisms, oil, and noticeable odour (e.g., sulphur, chlorine, sewage, petrol). Because odours from contaminated sediments could be potentially hazardous, care should be taken to avoid smelling them closely and directly.

- the intact core samples (liners) should be secured in an upright position (e.g., rack) and labels should be applied with the appropriate information (sampling site location, sample number, and/or identification, time and date of collection, method of collection, name or initials of the collector) to ensure accurate sample identification.

- air should be excluded from the liners and the volume of overlying water should be minimised to reduce the potential for resuspension of the surface sediments during transport. Particular care should be taken to retain the surficial floc overlying the core. The core tubes should be placed into either a transport container in an upright position (e.g., insulated box or cooler) with ice packs or into a refrigerated unit that can maintain a preferred temperature of 4 ± 2 °C, which can be continuously monitored during transport.

- a travel blank should be a part of the QA/QC program.

- subsampling, if necessary, should be performed using extrusion procedures within 24 hours of collection and should be restricted to parts of the sediment sample which have had no direct contact with the sampler.

- box-core samples should be subsampled using clean, non-reactive (e.g., Teflon®-lined scoop or hand-coring device) implements for the top 0-5 cm and a hand-corer for greater depths; subsampling should be restricted to sediments which have had no direct contact with the sampler.

- if compositing of samples or subsamples is required, it can be done in the field or in the laboratory. Before a composite sample or subsample may be analysed or used in a test it should be homogenised with a mechanical mixer, or by hand, until uniform in colour and texture, or for a specified period of time that has been determined experimentally.

- if the maintenance of an O_2-free environment for sediment handling is part of the study plan, handling of anoxic sediments should be done in a glove box in the presence of an inert gas (e.g., nitrogen).

Table 8-5: Advantages and disadvantages of the different sediment corers most commonly used to collect sediments (ASTM 1992a; Mudroch and MacKnight 1991)

Device/Dimensions	Use	Sediment Depth Sampled (cm)	Volume of Sediment Sample (cm^3)	Advantages	Disadvantages
Fluorocarbon plastic or glass tube 3.5-7.5 cm I.D.; ≤120 cm long	Shallow wadeable waters or deep waters if SCUBA available. Soft or semi-consolidated deposits	0-10	96-442	Preserves layering and permits historical study of sediment deposition; rapid; samples immediately ready for laboratory shipment; minimal risk of contamination	Small sample size requires repetitive sampling
Hand corer with removable fluorocarbon plastic or glass liners 3.5-7.5 cm I.D. ≤120 cm long	Same as above except more consolidated sediments can be obtained	0-10	96 - 442	Handles provide for greater ease of substrate penetration; above advantages apply	Careful handling necessary to prevent spillage; requires removal of liners before repetitive sampling; slight risk of metal contamination from barrel and core cutter
Box corer	Same as above but the depth of the unconsolidated sediment should be at least 1 m	0-50	≤ 30,000	Collection of large sample, undisturbed, allowing for subsampling	Hard to handle; heavy machinery required
Gravity corers: Phleger Corer 3.5 cm I.D. ≤50 cm long	Deep lakes and rivers. Semi-consolidated sediments	0-50	≤481	Low risk of sample contamination; maintains sediment integrity relatively well; high point loading with sharp cutting edge	Careful handling necessary to avoid sediment spillage; small sample requires repetitive operation and removal of liners. Time consuming
Kajak-Brinkhurst Corer 5 cm I.D. ≤ 70 cm long	Deep lakes and rivers. Soft fine-grained sediments	0-70	≤ 1374	Greater volume than the Phleger Corer	Same as Phleger Corer
Benthos Gravity Corer 6.6, 7.1 cm I.D. ≤ 3 m long	Soft, fine-grained sediments	0-3 m	≤ 10263	No loss of sample from tube because the valve is fitted to the core liner; fins promote vertical penetration	Weights required for deep penetration so the required lifting capacity is 750 to 1,000 kg; vertical penetration is required; sediment compaction

Table 8.5: Advantages and disadvantages of the different sediment corers most commonly used to collect sediments (ASTM 1992a; Mudroch and MacKnight 1991) (continued)

Device/Dimensions	Use	Sediment Depth Sampled (cm)	Volume of Sediment Sample (cm^3)	Advantages	Disadvantages
Alpine Gravity Corer 3.5 cm I.D.	Soft, fine-grained, semi-consolidated substrates	≤ 2 m	≤ 1924	Interchangeable steel barrel for different penetration depths	Requires a vertical penetration but has no stabilising fines; penetration is often non-vertical and incomplete and requires a lifting capacity of 2,000 kg; sheared laminae and disturbed sediment integrity; compaction of sediment
Piston Corers	Ocean floor and large deep lakes. Most substrates	3 to 20 m	?	Typically recovers a relatively undisturbed sediment core in deep waters	Requires lifting capacity of >2,000 kg; non-activation of piston and piston positioning at penetration; disturbance of the surface (0-0.5 m) layer
BMH-53 Piston Corer	Waters of #2m ft deep when used with extension rod. Soft to semi-consolidated deposits	> 2 m	?	Piston provides for greater sample retention	Cores should be extruded onsite to other containers; metal barrels introduce risk of metal contamination
Boomerang Corer 6.7 cm I.D.	Ocean floor (up to 9,000 m deep)	1 m	3525	Minimal shipboard equipment so small vessels could be used	Only 1.2 m penetration; recovery requires calm water; 10-20% loss rate
Vibratory Corer 5-7.5 cm I.D.	Continental shelf of oceans, large lakes. Sand, silty sand, gravelly sand substrates	3-6 m	5,890-13,253	For deep profiles it effectively samples sandy substrates with minimum disturbance; can be operated from small vessels (e.g., 10 m long)	Labour intensive; assembly and disassembly might require divers; disturbance of the surface (0-0.5 m) layer

8.4.3.2 Transport and Storage of Core Samples

If sediment core samples are not sectioned or subsampled in the field, they may be stored temporarily in the field before they are transported intact, in the core liner, directly to the laboratory. If sectioning or subsampling takes place in the field, then the subsamples or sections may also be transferred to sample containers and stored temporarily. The sample containers with the field-collected sediments are then placed into a transport container and shipped to the laboratory. If the transport container cannot accommodate the dimensions of the long core samples (core liners > 1 m), then the core samples can be cut prior to transport into 1 m lengths, and the ends securely capped such that there is no air trapped inside the liners. It should be demonstrated that the method and equipment used to cut or section the core causes minimal disruption to the integrity of the sediment within the section, does not contaminate the sample, and results in no loss of the sediment continuum. A variety of methods have been used to cut, split, or section sediment cores (Mudie et al. 1984; see Mudroch and MacKnight 1991, for review).

Freezing prior to transport or sectioning is not recommended because freezing can change the sediment volume, which varies with water content, and it permanently changes the structure of the sediment (de Groot and Zschuppe 1981; Rutledge and Fleeger 1988; Mudroch and MacKnight 1991). While in transit to a temporary storage facility or to the laboratory, frozen samples should not be thawed, and the temperature of samples that are temporarily in cold storage (e.g., 4 ± 2°C) should not exceed 7°C nor fall below 1°C.

8.4.3.3 Manipulation of Collected Samples

A major goal associated with the collection and storage of sediment cores is to maintain the integrity of the samples until their physical and chemical characteristics and/or toxicity are evaluated. This is accomplished by minimising changes to the physical, chemical and biological properties of the samples and subsamples that might result from transportation and storage. All manipulations and procedures used to prepare the test samples can potentially alter the nature of the field samples. Manipulation procedures, therefore, are intended to prepare or process field samples so that they remain as representative as possible of the sediment at the site of collection in terms of chemistry and toxicity, while making the sample more amenable for biological testing or characterization.

Sample preparation may involve physical operations such as sieving, drying, freeze-drying, crushing, grinding and/or chemical treatment steps such as dissolution, extraction, digestion, fractionation, and derivatisation of some of the organic contaminants, as well as pH adjustment, and the addition of reagents (Keith 1992, Section 8.1.7.6). These sample preparations, although necessary, are also potential sources of bias, variance, contamination, chemical alteration, and sediment loss. Therefore, sample preparation should be included in the study plan and fully documented to give a complete history of the transformation from a field-collected sediment to a test sediment. For Investigation of

Cause Monitoring study techniques such as toxicity testing, preparation of anoxic test sediments should be performed in a glove box in the presence of a controlled flow of an inert gas, if it is desirable to maintain these anoxic conditions.

8.4.4 Sediment Chemistry

8.4.4.1 Partial Metal Concentrations in Sediment (Partial Extraction), Sequentially Extracted Metals, SEM/AVS Ratios

Although total metal concentrations may not be directly related to biological availability and toxicity, many sediment quality guidelines are currently based on total metal concentrations. A variety of methods have been used to predict the biological effects of metals from metal contaminated sediments. These include the normalisation of sediments for particle size, organic content, or extractable fraction of metals using AVS (Acid Volatile Sulphides) and SEM (Simultaneously Extracted Metals).

It is generally thought that a particular chemical form of an element determines its behaviour, biological availability and potential toxicity, rather than the total concentration in sediments. Specific chemical forms can be measured:
a) by direct instrument techniques;
b) directly by sequential digestion of sediments; or can be predicted by
c) predicting levels through thermodynamic modelling.

Direct instrument techniques include: X-ray photoelectron spectrometry (XPS); scanning electron microscopy/X-ray microanalysis (SEM/XRM); secondary ion mass spectrometry (SIMS); and, Auger electron spectrometry (AES). These methods have been applied to geochemical studies and for mineral exploration.

The relative strength of association between metals and particles can be assessed by single or sequential extraction or sediment digestion methods. Weak acids or chelating agents (e.g. EDTA) and reducing agents may be used to differentiate between different chemical forms. Sediment fractions can be operationally defined (e.g., ferromanganese oxyhydrides) depending on the digestion method used. The recent acid volatile sulphide (AVS) concept assumes that metal concentrations in pore water of anoxic sediments are controlled by sulphides. AVS are extracted by the cold-acid purge and trap technique. Simultaneously extracted metals (SEM) represent the portion of total metals released during AVS dissolution. The SEM/AVS ratio is sometimes used to characterise metal availability. When the SEM fraction exceeds the AVS fraction (e.g. SEM:AVS ratio >1) then the free metal may be present in the pore water at levels adequate to cause acute toxicity.

The limitation of these analyses is that anoxic sediment samples must be carefully collected and stored to prevent oxidation. A description of sediment collection methods can be found in Warren et al. (1998). Partial extraction and AVS procedures are not routine for most labs and are relatively expensive. During the AETE field program no relationship was identified between sediment toxicity and the SEM/AVS ratio.

8.4.4.2 C:N Ratio for Marine Sediment

Effects on the benthic invertebrate community may occur as a result of organic enrichment in sediments. To determine if organic enrichment is contributing to effects, a combination of measurement techniques should be used in the marine environment. The measurement of total organic carbon (TOC) provides an indication of organic enrichment. Measuring sediment Eh (redox) provides an indication of the oxygen conditions in marine sediments. Measuring sulphides in the sediment provides an indication as to whether the breakdown of organic sediment material is occurring (Environment Canada 1994).

Measuring the carbon to nitrogen ratio (C:N ratio) in marine sediments should provide an indication of the source of the organic enrichment. If the organic enrichment is a result of land based sources (e.g. municipal sewage, pulp and paper effluent), the C:N ratio will be higher (Hargrave et al.1995). If the organic enrichment is a result of the natural source such as the breakdown of marine aquatic plants, the C:N ratio will be much lower. Therefore, if an benthic invertebrate community study conducted indicates an effect on the benthic invertebrate community, and there is evidence that the effect could be due to organic enrichment (elevated TOC, elevated Eh), determining the marine sediment C:N ratios can help identify the source of the organic loading to that ecosystem.

8.4.5 Whole Sediment Toxicity Testing

Sediment toxicity tests may be used to evaluate potential contamination in aquatic environments. These tests provide a direct method to determine chemical availability of contaminants in sediment, and can be used as an evaluation tool in conjunction with chemical data. Whole sediment toxicity tests provide direct measurement of toxicity by evaluating the effects of sediment on test organisms. Test results are expressed as "toxic" in comparison with a reference site. Effects can also be expressed as the difference (in %) with response in control or reference sediments.

Whole sediment toxicity testing can be used to verify that alterations seen in benthic invertebrate communities are due to toxicity of sediments and not other physical or biological factors. Adverse effects on benthic invertebrate community structure could be due to sediment toxicity, but could also be due to other factors (predation, habitat differences, etc.). Concurrent impairment of benthic invertebrate community structure and sediment toxicity implicates the sediment itself as the cause of effects on the benthic invertebrate community. These tests also provide important information for interpretation of field effects in situations where benthic invertebrate community data are inconclusive, or if only pollution tolerant species are present in both exposure and reference areas.

As well, the use of whole sediment toxicity, sediment contaminant analysis, and benthic invertebrate community studies within a common sediment depositional area allows for data correlations between the different sediment monitoring tools, through the application of the Sediment Quality Triad (see Section 8.4.5). Strong correlations indicate that contaminants in the sediment are involved in observed field effects.

The limitation of whole sediment toxicity tests is that they are laboratory tests that do not determine *in situ* toxicity directly. Further, if a standard overlying water is used (as is commonly done), instead of water collected from the study site, then the effects of differences in overlying water quality on sediment toxicity, if present, are not taken into account.

For all test types, detailed records of all aspects of the samples, test organisms, test conditions, equipment and test results are kept by the laboratory. Control exposures with clean sediment are performed at the same time as sample tests. Testing of reference toxicants is done in "water only" exposures and the data plotted on a control chart. If individual results are within expected limits, the performance of the batch of test organisms is ensured. Test requirements include a temperature controlled room, bioassay equipment (exposure containers), invertebrate culture facilities and a source of clean non-chlorinated water. Technical personnel should be skilled in invertebrate biology and culture and in bioassays.

8.4.5.1 Whole Sediment Toxicity Test: Survival and Growth of *Hyalella azteca*

The *Hyalella azteca* test is frequently observed to be one of the most sensitive in studies of metal impacted sites (Beak and Golder 1998). Standard test procedures are commercially available across Canada (Environment Canada 1997). Comparison of growth and survival endpoints in the test can sometimes be useful in cause identification. For example, growth is affected at lower concentrations of copper in the sediment than is survival, but the reverse is true for zinc and lead (Borgmann and Norwood 1997, 1999).

Fifteen individuals (aged 1-10 days) are exposed to sediment samples and bioassay water. Mortalities are monitored daily throughout the test. After 14 days, the sediment is sieved and the surviving animals are counted, dried and weighed. Endpoints are survival and average growth. For QA/QC, detailed records of all aspects of the samples, test organisms, test conditions, equipment and test results are kept by the laboratory. Control exposures with clean sediment are performed at the same time as sample tests. Testing of reference toxicants is done in "water only" exposures and the data plotted on a control chart. If individual results are within expected limits, the performance of the batch of test organisms is ensured.

The main limitation of this test is that *Hyalella azteca* is not a tolerant species. Under lethal conditions for the organism, growth becomes irrelevant and cannot be used to measure relative differences in toxicity.

8.4.5.2 Whole Sediment Toxicity Test: Survival and growth of the freshwater midges Chironomus tentans or Chironomus riparius

Chironomids are often used to test sediments in the US and Canada. The test method is described in Environment Canada (Environment Canada 1997): Test for Growth and Survival in Sediment Using Larvae of Freshwater Midges (*Chironomus tentans* or *Chironomus riparius*).

Ten second (*C. tentans)* or first *(C. riparius)* instar organisms are exposed to a sediment sample and bioassay water. Losses of overlying water due to evaporation are replaced during the test. After 10 days, the sediment is sieved and the surviving animals are counted, dried and weighed. Endpoints are survival and average growth. For a valid test, there must be >70% survival in the control sediment and mean growth must be ≥0.6 g (*C. tentans*) or ≥0.2 g (*C. riparius*).

8.4.5.3 Whole Sediment Toxicity Test: Survival and reproduction of the oligochaete *Tubifex tubifex*

The test is described in the Ontario Ministry of the Environment "Laboratory Sediment Biological Testing Protocol" (Bedard et al. 1992).

Tubifex individuals are exposed to equal amounts of sediment and bioassay water. Sexually mature individuals (aged 8 weeks) are introduced and incubated for 28 days. The production of cocoons indicates reproduction of the organisms. At the end of the test, the sample is sieved and the number of surviving adults, the number of full and empty cocoons, the number of young less than 500 mm and the number of young greater than 500 mm are counted as measurements of survival and reproduction. This test is less sensitive than the *Hyalella azteca* test.

8.4.5.4 Whole Sediment Toxicity Test: Survival and growth of the mayfly *Hexagenia limbata*

The test is described in the Ontario Ministry of the Environment "Laboratory Sediment Biological Testing Protocol" (Bedard et al. 1992).

The test uses 3 to 4 month old early instar *Hexagenia limbata* mayfly nymphs which are laboratory-reared from field-collected eggs. Ten to fifteen *Hexagenia limbata* individuals are exposed for 21 days to bioassay water and sediment. Animals are not fed during this time. At the end of the test, the number of surviving animals are counted and the animals are dried and weighed to determine dry weight and to give an indication of growth. Survival and growth in the test sediment are compared statistically to survival and growth in the control.

This test is sensitive to metals, and is ecologically relevant since mayflies occur across Canada.

8.4.5.5 Using a battery of whole sediment toxicity tests: *Hyalella azteca*, *Chironomus tentans* or *Chironomus riparius*, *Tubifex tubifex* and *Hexagenia limbata*

Batteries of tests can be used to demonstrate sediment toxicity in more than a single species. Use of a battery of tests instead of just one test (e.g. the *Hyalella azteca* test) helps to demonstrate the significance of the sediment toxicity response (i.e. number of test species responding). This provides additional information on the magnitude and extent of sediment toxicity. The relative responses observed in sediment tests with different species can aid in the identification of the cause of sediment toxicity if information is available in the literature on the relative sensitivity of different species to different metals. Multi-species testing can also aid in explaining the relationship between benthic invertebrate community structure and sediment toxicity. Use of multiple species in toxicity tests allows for more direct comparisons of abundance/absence of benthic invertebrate species with toxicity test results for the same, or closely related species. Standard test procedures are available (Bedard et al. 1992, ASTM 1997, Environment Canada 1997).

8.4.5.6 Metals in Overlying Water in Whole Sediment Toxicity Tests

Measurement of metals in overlying water in whole sediment toxicity tests can be used for the quantification of bioavailable metal or the identification of cause of sediment toxicity.

Measurement of dissolved metal in the overlying water can provide information on the relative bioavailability, and hence potential toxicity, of metals in sediments, because the toxicity of metals in sediments is not proportional to total metal concentrations in sediment, since bioavailability can vary greatly from one sediment to another. Metal bioavailability, especially to *Hyalella azteca*, often appears to be primarily due to dissolved metal in the overlying water (Deaver and Rodgers 1996, Warren et al. 1998, Borgmann 1999, Borgmann and Norwood 1999). If data on metal toxicity in water are available, and if toxicity is known to be primarily due to dissolved metals for the benthic invertebrate species in question, then water concentrations can sometimes be used to infer the cause of toxicity. Use of metal concentrations in overlying water is restricted to static toxicity tests in which the overlying water is not renewed.

The main limitation of this type of measurement is that the toxicity of metals in water can be reduced substantially by complexation with dissolved organic matter (DOC) leaching from sediments. Water concentrations resulting in a given level of toxicity may, therefore, be higher in sediment toxicity tests than in water-only tests (Borgmann 1999). This effect can be reduced by using a large water to sediment ratio in the toxicity test (Borgmann and Norwood 1999). Water chemistry other than DOC may also affect metal toxicity, and metal concentrations in overlying water should be compared to metal concentrations causing toxicity in water-only tests only if the major ion composition of the water used in these tests is similar. This technique will not help identify the cause of toxicity if the solid phase of the sediment contributes significantly to toxicity.

8.4.5.7 Bioaccumulation: Measurement of Metals in Test Organisms from Sediment Toxicity Testing

Measurement of metals in test organisms from sediment toxicity testing can be used for the quantification of bioavailable metal, and the identification of cause of sediment toxicity. Metal concentrations in benthic invertebrates exposed to sediments are a direct measure of bioavailability and correlate much better with metal-induced sediment toxicity than total metals in the sediment (Borgmann and Norwood 1997a, 1997b). If data on the relationship between bioaccumulation and toxicity are available, metal bioaccumulation can be used directly to infer cause of toxicity (Borgmann and Norwood 1997b).

Animals for bioaccumulation measurement can be obtained from the same tests used to measure sediment toxicity. A gut clearance period is normally recommended (Neumann et al. 1999). If chronic toxicity tests result in complete mortality, tests can be repeated using older animals and a shorter (e.g., 1 week) exposure period (i.e., to reduce toxicity). Body concentrations of metals in *Hyalella azteca* equilibrate within about 1 week (e.g., MacLean et al. 1996, Borgmann et al. 1997a). If data on the relationship between bioaccumulation and toxicity are not available, these data may be obtained using sediment spiking experiments as part of an identification of cause study (e.g., Borgmann and Norwood 1997a, 1999).
This method is particularly useful for identification of bioavailability of non-essential metals (Cd, Pb, Tl, etc.). However, copper and zinc in invertebrate tissues may be regulated to varying degrees, and their background concentrations may be high. This may limit the usefulness of copper and zinc measurements in invertebrates.

8.4.6 The Sediment Quality Triad

The use of the sediment quality triad approach is strongly recommended during Investigation of Cause of effects on benthic invertebrate communities. This approach, or elements of this approach, may also be used during Focussed Monitoring of benthic invertebrate communities, to aid in the interpretation of results. AETE endorsed the use of the sediment quality triad, for use on a site-specific basis, but not for "routine" monitoring (ESG 1999).

The usefulness of sediment quality criteria (Environment Canada 1994) alone is limited due to the following:

a) single chemical concentrations do not integrate effects;
b) the criteria do not take into account site-specific conditions (i.e. they are based on averages from many diverse sites); and,
c) the provincial criteria are based on Great Lakes sediments and may not reflect stream, riverine or headwater conditions.

In other words, exceedence of these criteria at a site does not necessarily mean that there are environmental effects at that site as a result of contaminated sediments. Hence a site-specific sediment quality triad that integrates chemical, biological and ecological effects is more meaningful (Chapman 1992). The triad integrates sediment chemistry, toxicity and resident biota (the community structure of benthic invertebrates), making the data more powerful than if each component were interpreted individually.

The three corners or the triad can be designated as chemical [C], biological [B], and toxicity [T]. Results from each "arm" of the triad can be statistically examined to determine correlations, if any. The results from the various components of the SQT can be integrated into a weight-of-evidence analysis which can be presented both qualitatively and quantitatively in a number of different interpretation and presentation formats (see Alden 1992; Chapman 1992; Warwick and Clarke 1991).

Overview of the Sediment Quality Triad study approach:

1) collect sediment samples
2) analyse sediments for chemical parameters
3) conduct laboratory toxicity tests on sediments
4) identification and enumeration of benthic invertebrates in sediment samples
5) collect fish for tissue analysis (optional)
6) data analysis and report preparation

1) Sample Collection

Samples are collected from both reference and exposure areas or over an exposure gradient to help delineate the geographic extent of the effect. The total number of sampling locations should be at least ten. Ten sampling locations should provide enough data for a robust statistical analysis. Samples at each location should be collected for chemistry, benthos and sediment toxicity.

2) Sediment Chemistry

Chemical analysis of the sediments should include particle size and total organic content. Any or all of the parameters listed in Table 8.4 should be analysed. The decision should be based on previous information from the site particularly from water and effluent chemistry.

3) Sediment Toxicity Tests

A battery of three tests is ideal for the triad approach, involving three different test species (MOEE 1997; ESG 1999). This provides a safeguard that a biological effect is not missed if one test species is not sensitive to the chemicals present in the sediments. It also provides test organisms from different trophic levels in the ecosystem so that the major relevant trophic groups are represented. The recommended tests use benthic invertebrates and fish, and incorporate different biological parameters (i.e., survival and growth).

The three recommended tests are:

- *Hyalella azteca* 14-day growth and survival (Environment Canada 1998);
- *Chironomus tentans* 14-day growth and survival (Environment Canada 1998); and,
- fathead minnow 21-day toxicity and bioaccumulation (MOEE 1993).

Juvenile fathead minnows used in the 21-day test can be preserved for future whole body residue analysis to determine bioaccumulation of chemicals from the sediments. These toxicity and bioaccumulation results would provide data comparable to the MOEE (1997) study. Although a battery of tests is ideal, it is possible to start with just *Hyalella azteca*.

4) Identification and enumeration of benthic invertebrates in sediment samples

The identification and enumeration of benthic invertebrates should be conducted as detailed in Chapter 5 of this guidance document.

5) Fish tissue collection and analysis (optional – consider in areas where there is a concern about the human consumption of fish)

The collection of fish for metals analyses would provide information on whether or not chemicals in resident biota may be available for human consumption.

6) Data analysis and Report Preparation

Data obtained from the benthic survey, chemical analysis and toxicity tests will be statistically analysed and integrated in a weight of evidence approach.

8.4.7 Sediment Pore Water Monitoring

Fine-grained surface sediments in lakes typically contain 90 to 95% water (Adams 1991). Some of this water is bound to the crystalline lattice of minerals in the sediments, but most of the water simply occupies the space between sediment particles. This water is referred to as pore or interstitial water. The intimate association of this water with the surface of sediment particles results in reactions between the particles and the water that approach equilibrium. The partitioning of contaminants in sediments between the particulate and water phases depends to a large extent on the amount of organic carbon, sediment particle size, the chemical form of the contaminants, and the physiochemical environment (e.g., pH, temperature, redox potential, sorption/desorption properties of sediments, or the equilibrium between the solid and liquid phases). The dynamics of these processes are not well understood; however, it is generally assumed that concentrations of most substances in the pore water approach equilibrium with the solid phase and its associated contaminants, and that metals in pore water largely represent the biologically available fraction in sediments. Consequently, pore water has been collected for toxicity testing to approximate the relative toxicity of contaminated sediment, and/or to assess contaminant levels.

The nature of sediments at the study site can largely influence the usefulness of pore water measurements. Sediments which are either very coarse grained or hard, compacted clays, will not likely have pore waters that are significantly contaminated. Therefore, sampling of pore waters should be restricted to sediments ranging from sandy to non-compacted clays.

If sediments are anoxic, as most depositional sediments are below 2 cm in depth, then all steps involved in sample collection and processing should be conducted in an inert atmosphere or with limited exposure to prevent oxidation and subsequent sorption/precipitation of reduced metal species if metal speciation is of interest. When anoxic sediments are exposed to air, volatile sulphides may also be lost, which may increase the availability (and toxicity) of sulphide-bound metals. Finally, pore water samples undergo rapid chemical changes giving a storage life of only hours to days.

There is a relatively large amount of literature describing toxicity testing with pore water (Burton 1998). However, it is not as extensive as whole sediments, and there are few standardised methods for toxicity testing of freshwater organisms with pore water. Environmental Canada (1992) has two standard methods using pore water for Echinoids (Sea Urchins and Sand Dollars) and luminescent bacteria.

8.4.7.1 Collection of Pore Water

Sediment pore water can be isolated using either laboratory or field (*in situ*) approaches. The principle laboratory methods for collection of pore water from sediment are centrifugation, pressurisation, and suction. Laboratory methods allow for extraction of greater volumes of water, and are preferred when samples are being collected for toxicity testing. Note that samples for toxicity testing have a very short storage time.

Field collection using "peepers" is the most accurate method to obtain representative samples and is recommended for geochemical investigations. Peepers are small diffusional chambers with membrane or mesh walls, filled with site water, gels or non-polar solvents which are buried in the sediments and allowed to equilibrate with the surrounding pore water. The chambers are left in place for 2 to 20 days. Peepers generally only collect very small volumes of water, and as a result the minimum volume of sample required for chemical analyses should be confirmed by the analytical laboratory prior to sampling. Chemical analyses of pore water samples should be conducted following methods described in Chapter 6 for water and effluent.

1) Field (*in-situ*) Collection of Pore Water

The *in-situ* sampling of pore water has been achieved by various methods (Barnes 1973; Sayles et al. 1973; 1976; Hesslein 1976; Mayer 1976; Murray and Grundmanis 1980; Brinkman et al. 1982; Whiticar 1982; Bottomley and Bayly 1984; Belzile et al. 1989; Buddensiek et al. 1990; Carigan and Lean 1991). There are too few comparative studies to recommend one method over another. Most *in-situ* pore water samplers have been designed to sample pore water close to the sediment-water interface by either diffusion through a membrane sandwiched between acrylic plates (e.g., dialyzer or peeper), or by suction and filtering. These methods have evolved in response to the realisation that separating pore water from collected sediment in the laboratory produces results that do not always represent the nature or chemistry of the water (Kriukov and Manheim 1982; Howes et al. 1985; Adams 1991). These inconsistencies result primarily from changes in temperature (Bischoff et al. 1970), pressure (e.g., CO_2 degassing) (Simon et al. 1985), and the exposure of both the liquid and particulate phases of sediment to oxygen (Bray et al. 1973; Sayles et al. 1973; Hesslein 1976; Mayer 1976; Lyons et al. 1979; Lee and Jones 1984).

The various methods for collection of *in-situ* pore water and their respective attributes and limitations are summarised in Table 8.6. For more details consult Environment Canada (1994) and/or the references therein. It is important to be aware that all of the techniques use some type of filter, membrane, or plastic material such that sorption of organic contaminants can become a significant problem. Additionally, these techniques usually generate only small volumes of pore water and so both chemical analysis and biological testing for organic contaminants might be problematic. More details on *in-situ* collection of pore water can be found in Appendix 8.1.

2) Laboratory Recovery of Pore Water from Bulk or Core Sediment Samples

Manipulations for Extraction of Pore Water

Squeezing of sediments to extract pore water can be done by high or low pressure mechanical squeezing, or low-pressure gas squeezing. The squeezing devices generally consist of a series of clamps, filters and membranes in conjunction with collection devices for retaining pore water. The composition of the materials used in the construction of a squeezing device varies and may include Teflon®, stainless steel, brass and nylon. A comprehensive description of a large number of squeezers including the composition of their parts, the pressure used, the sample volume processed, and the volume of pore water obtained is provided in Adams (1991).

Factors that can influence the composition of pore water obtained using the squeezing technique include contamination from component materials of the squeezing device, operational conditions, and chemical fractionation of the sample. Details on these factors can be found in Appendix 8.2.

Table 8-6: A summary of the various methods for the collection of pore water, *in-situ*

Method of Collection	Advantages	Disadvantages	Reference
Long-needled syringes	closed system minimises the possibility of contamination	restricted to use in shallow waters (e.g., <10 m); small volumes (5 to 10 mL) can be collected from a given depth	Bauer et al. 1988 Brinkman et al. 1982
Modified syringes	absence of contamination from overlying water	effectively sampled <350 μL	Knezovich and Harrison 1987
Probe system	repeated sampling at specific depths (0 to 35 cm; 1-cm increments); low cost, easily constructed; stable unit which could be operated throughout the year	requires a 2-week period for sediment stabilization; potential problem with mixing within adjoining horizons; requires adjustments of the zero level because of shifting sediment surfaces	Buddensiek et al. 1990
Hydropressure or vacuum filtration	closed system minimises the possibility of contamination; can be used in shallow (<10 m) or deep waters (20 to 1800 m); displacement with inert gas maintains anoxic conditions and minimises artefacts; can be used for estuarine sediments, rivers, lakes and marshes, deeper marine waters or shallow subtidal marine waters	although volumes were slightly larger (50 to 75 mL) the depth from which the interstitial water is retrieved is uncertain (e.g., limited depth resolution)	Montgomery et al. 1979, 1981; Hertkorn-Obst et al. 1982; Howes et al. 1985; Sayles et al. 1976; van der Loeff 1980
Dialysers or peepers	use of specific membranes can confer some solute selectivity; relatively free of temperature, pressure and oxidation artefacts; simple and less costly to construct and operate than most other methods	cellulose-based membranes are susceptible to microbial attack; equilibrator and system should be degassed with nitrogen to ensure deoxygenation; equilibration time should be determined experimentally as it is variable and may be lengthy; oxidation of interstitial water during careless transfer from dialyser; development of electrical potential across membrane	Adams 1991 Martens and Klump 1980 Hesslein 1976 Kelly et al. 1984 Carignan and Tessier 1985 Tessier et al. 1989 Simon et al. 1985 Carignan 1984

8.4.7.2 Sediment Pore Water Toxicity Testing

The concentrations of contaminants in pore water may be more highly correlated with toxicity to aquatic organisms (Cairns et al. 1984; Nebeker et al. 1984; Schuytema et al. 1984; Knezovich and Harrison 1988; Giesy and Hoke 1990) than those in bulk sediments (Patrick et al. 1977; Adams et al. 1985; Shaner and Knight 1985; van de Guchte and Mass-Diepeveen 1988; Di Toro 1989; Ankley et al. 1991c; Carr and Chapman 1992). Therefore, toxicity testing using sediment pore water can be an important complement to whole sediment toxicity testing. The relative importance of pore water and whole sediments as sources of toxic contaminants to aquatic organisms appears to depend on the species of test organism and the type of contaminant (Knezovich and Harrison 1987; Giesy and Hoke 1990; Harkey et al. 1994).

Tests which may be used for pore water toxicity testing are briefly described below. Note that these are standard tests for toxicity testing of water and effluent, and are not unique to pore water testing.

1) Sea Urchin Fertilisation Bioassay

Sea urchin fertilisation test is described in the chapter on sublethal toxicity, Section 7.2.

2) Microtox Acute Test

The Microtox test is used to evaluate the effect of effluent exposure on light production by the naturally luminescent marine bacteria, *Vibrio fischeri.* The result is expressed as the concentration where light output is reduced by 25% or 50% (IC25, IC50).

The Microtox test is frequently used on site by many industries. The test is described in the manufacturer's handbook "Microtox Manual. A Toxicity Testing Handbook." (Microbics Corp. 1992). The Microtox test is a rapid screening bioassay kit, which measures toxic effects on the light output of a standardised luminescent bacterial culture. The light output of the bacteria is measured after 5 or 15 minutes exposure to the sample. This is a rapid assay since the duration of the entire test is only 30 minutes. For QA/QC, detailed records of all aspects of the samples, test organisms, test conditions, equipment and test results are validated and kept by the laboratory. Reference toxicant tests are performed at the same time as samples are tested to validate the performance of the batch of test organisms.

The main limitation of this method is that the test organism is a marine bacteria, with little relevance to Canadian mining environments.

3) Acute Lethality with *Daphnia magna*

The purpose of this test is to determine the concentration of test water that causes 50% mortality to *Daphnia magna* during a 48 hour exposure period. This standard tests involves placing groups of less than 24 hour old *D. magna* neonates into a range of test water concentrations. Test water is diluted with freshwater to which the invertebrates have been acclimated. Toxicity tests are conducted in test tubes, generally 55 mL in size. For each concentration, including controls, 4 replicate tubes are set up, each with 3 daphnids, for a total of 12 individuals per concentration. Tests are conducted at 20 ± 1°C. Both temperature and photo-period are similar to culture conditions. Tests are conducted under static conditions with no renewal of the test solutions. Observations for immobility or mortality are recorded at 24 and 48 hours. A daphnid is considered dead if there is no visible heartbeat upon microscopic examination. The *Daphnia magna* lethality test is described in detail in the Federal Protocol (Environment Canada 1990).

This test is widely used and accepted as a regulatory test throughout Canada.

4) Reproduction and survival using the cladoceran *Ceriodaphnia dubia.*

This test is described in the chapter on sublethal toxicity, Section 7.2.

5) Growth inhibition of the alga *Selenastrum capricornutum*

This test is described in the chapter on sublethal toxicity, Section 7.2.

8.4.8 Toxicity Identification Evaluation and Toxicity Reduction Evaluation (TIE/TRE)

If toxicity testing results indicate that either sediment or pore water is toxic, the mine may undertake further studies to determine and eliminate the cause(s) of toxicity. Protocols to investigate the probable causes of toxicity have been developed by the USEPA (USEPA 1989) and are known as Toxicity Reduction Evaluation (TRE) and Toxicity Identification Evaluation (TIE) studies. The use of TIE/TREs are relatively new in Canada and their application to mining was reviewed under the AETE program (ESG 1999).

The general objectives of TIE/TRE process are to (US EPA 1989):

a) evaluate the potential sources of toxicity;
b) characterise the toxicity observed in the sample;
c) provide a preliminary identification of the possible sources of this toxicity by evaluating changes that occurred in the toxicity following a variety of chemical and physical manipulations and treatments; and
d) ultimately provide measures for reduction and elimination of the toxicants.

A successful TIE/TRE will involve the co-ordination of a multidisciplinary team including toxicologists, chemists, engineers and very importantly, mine personnel.

The TIE portion of the TRE program is divided into three phases:

- Phase I - characterization of toxicity through a variety of treatments (USEPA 1991a);
- Phase II - identification of the suspected toxicants (USEPA 1993a); and
- Phase III - confirmation of suspected toxicants (USEPA 1993b).

Phase I TIEs are considered the next phase of assessment when sediments or pore water are identified as toxic in the USA.

The Phase I TIE methods were originally developed for use with acute lethality tests using fathead minnows or *Ceriodaphnia dubia* but have been adapted for sublethal testing (USEPA 1991b,c), and may also be used for pore water. In Canada, rainbow trout and *Daphnia magna* are commonly used during TIE tests.

The standard USEPA Phase I effluent characterization treatments involve filtration at different pH, aeration at different pH, C18 solid phase extraction at different pH, treatment with ethylenediaminetetraacetate (EDTA) and sodium thiosulphate. A significant portion of toxicity observed in industrial effluents is often attributed to pH effects. Therefore, pH adjustment is used throughout Phase I to provide more information on the nature of the toxicants. Other treatments include passing samples of effluent through columns with carbon or zeolite. In all cases, the toxicity of treated samples is compared to non-treated samples to determine which approach, if any, reduced toxicity.

At present, there is not an established protocol for conducting TIE on sediments, although efforts are underway to develop such a protocol, particularly in the United States. A team of USEPA researchers recently published a paper documenting one possible method (Burgess et al. 2000).

One of the most important benefits of the TIE/TRE process is that it incorporates the responses of organisms into the assessment of complex mixtures to determine the identity of the substance(s) responsible for toxicity. Attempts to use chemical screening alone to identify substances responsible for toxicity are typically unsuccessful.

The application and benefits from the TIE/TRE process cannot be realised unless a detailed toxicity study is completed. Determining the cause of transient or non-persistent toxicity can be difficult and may require the testing and analysis of a large number of samples.

It should be emphasised that the TIE approach is not standardised beyond Phase I and subsequent studies to identify the specific toxicants require experienced personnel. There is no Canadian protocol to conduct TIE/TREs. The TIE/TRE approach does not prove the cause of toxicity but rather uses a weight of evidence approach. Toxicants are also often identified on the absence of contrary evidence (Mount 1997). Even a complete TIE/TRE may not conclusively identify the source of toxicity. However, a full TIE/TRE approach can be useful to characterise, identify and reduce the sources and causes of toxicity in many cases.

8.5 References

Adams, DD. 1991. Sampling Sediment Pore Water, pp. 203-202, in: CRC Handbook of Techniques for Aquatic Sediments Sampling, Mudroch A and MacKnight SD (eds.) CRC Press, Inc. Boca Raton, FL, 210 p.

Adams DD, Darby DA, and Young RJ. 1980. Selected Analytical Techniques for Characterizing the Metal Chemistry and Geology of Fine-grained Sediments and Interstitial Water, p. 3, in: Contaminants and Sediments, Baker RA (ed.), Ann Arbor Science Publishers, MI.

Adams WU, Kimerle RA, and Mosher RG. 1985. Aquatic Safety Assessment of Chemicals Sorbed to Sediments, pp. 429-453, in: Aquatic Toxicology and Hazard Assessment, Cardwell RD, Purdy R and Bahner RC (eds.), American Society for Testing and Materials, Philadelphia, PA.

Alden, III, RW. 1992. Uncertainty and Sediment Quality Assessments: I. Confidence limits for the triad. Env. Tox. Chem. 11: 637-644.

Ankley GT, Schubauer-Berigan MK, Dierkes JR, and Lukasewycz MT. 1991. Sediment Toxicity Identification Evaluation: Phase I (Characterization), Phase II (Identification) and Phase III (Confirmation), Modifications of Effluent Procedures, National Effluent Toxicity Assessment Centre Technical Report, 68-91, Draft VII, 12 p.

Ankley GT, Schubauer-Berigan MK and Dierkes JR. 1991c. Predicting the Toxicity of Bulk Sediments to Aquatic Organisms with Aqueous Test Fractions: Pore Water Versus Elutriate, Environ. Toxicol. Chem., 10: 1359-1366 (1991c).

APHA. 1995. Standard Methods for Examination of Water and Wastewater. 19th edition, published by the American Public Health Association (APHA).

ASTM (American Society for Testing and Materials). 1992a. E 1391-90, Standard Guide for Collection, Storage, Characterization and Manipulation of Sediments for Toxicological Testing, pp. 1134-1153, in: 1992 Annual Book of ASTM Standards, Vol. II.04, Section 11, Philadelphia, PA.

ASTM (American Society for Testing and Materials). 1997. Standard test methods for measuring the toxicity of sediment-associated contaminants with fresh water invertebrates (Annex 3, *Hexagenia spp.*, and Annex 4, *Tubifex tubifex*). Volume 11.05, Standard E1706-95b.

Barnes RO. 1973. An *in-situ* Interstitial Water Sampler for Use in Unconsolidated Sediments, *Deep Sea Res.*, 20: 1125-1128.
Baudo R. 1990. Sediment Sampling, Mapping, and Data Analysis, pp.15-60, in: Sediments: Chemistry and Toxicity of In-place Pollutants, Baudo R, Giesy JP and Muntau H (eds.), Lewis Publishers, Inc., Chelsea, MI.

Baudo R, Giesy JP, and Muntau H. (eds.). 1990. Sediments: Chemistry and Toxicity of In-place Pollutants, Lewis Publishers, Inc., Chelsea, MI, 405 p.

Bauer JE, Montagna PA, Spies RB, Prieto MC and Hardin D. 1988. Microbial Biogeochemistry and Heterotrophy in Sediments of a Hydrocarbon Seep, Limnol. Oceanogr., 33: 1493-1513.

Beak International Inc. and Golder Associates Ltd. 1998. Summary and cost-effectiveness evaluation of aquatic effects monitoring technologies applied in the 1997 AETE field evaluation program. prepared for: Aquatic Effects Technology Evaluation Program, Natural Resources Canada, Ottawa, ON.

Bedard D, Hayton A, and Persaud D. 1992. Ontario Ministry of Environment laboratory sediment biological testing protocol. Water Resources Branch, Ontario Ministry of Environment, Toronto, ON, 26 pp.

Belzile N, De Vitre RR and Tessier A. 1989. *In-situ* Collection of Diagenetic Iron and Manganese Oxyhydroxides from Natural Sediments, Nature, 340 (6232): 376-377 (1989).

Bender MB, Martin W, Hess J, Sayles F, Ball L, and Lambert C. 1987. A Whole-core Squeezer for Interfacial Pore-water Sampling, Limnol. Oceanogr., 32(6):1214-1225.

Bischoff JL, Green RE and Luistro AO. 1970. Composition of Interstitial Waters of Marine Sediments: Temperature and Squeezing Effects, Science, 167: 1245-1246.

Blomqvist S. 1990. Sampling Performance of Ekman Grabs – *In-situ* Observations and Design Improvements, Hydrobiologia, 206: 245-254.

Borgmann U and Norwood WP. 1997. Toxicity and accumulation of zinc and copper in *Hyalella azteca* exposed to metal-spiked sediments. Can. J. Fish. Aquat. Sci. 54: 1046-1054.

Borgmann U and Norwood WP. 1997a. Toxicity and accumulation of zinc and copper in *Hyalella azteca* exposed to metal-spiked sediments. Can. J. Fish. Aquat. Sci. 54: 1046-1054.

Borgmann U and Norwood WP. 1997b. Identification of the toxic agent in metal-contaminated sediments from Manitouwadge Lake, Ontario, using toxicity-accumulation relationships in *Hyalella azteca*. Can. J. Fish. Aquat. Sci. 54: 1055-1063.

Borgmann U and Norwood WP. 1999. Assessing the toxicity of lead in sediments to *Hyalella azteca*: the significance of bioaccumulation and dissolved metal. Can. J. Fish. Aquat. Sci. in press.

Borgmann U. 1999. Methods for assessing the toxicological significance of metals in aquatic ecosystems: Bioaccumulation-toxicity relationships, water concentrations, and sediment spiking approaches. Aquat. Ecosyst. Health Management. in press.

Bottomley EZ and Bayly IL. 1984. A Sediment Pore Water Sampler Used in Root Zone Studies of the Submerged Marcophyte, *Myriophyllum spicatum*, Limnol. Oceanogr., 29: 671-673.

Bray JT, Bricker OP, and Troup BN. 1973. Phosphate in Interstitial Waters of Anoxic Sediments: Oxidation Effects During Sampling Procedures, Science, 180: 1362-1364.

Brinkman AG, van Raaphorst W and Lijklema L. 1982. *In-situ* Sampling of Interstitial Water from Lake Sediments, Hydrobiologia, 92: 659-663.

Buddensiek V, Engel H, Fleischauer-Rössing S, Olbrich S and Wächtler K. 1990. Studies on the Chemistry of Interstitial Water taken from Defined Horizons in the Fine Sediments of Bivalve Habitats in Several Northern German Lowland Waters, I: Sampling Techniques, Arch. Hydrobiol., 119(1): 55-64.

Burgess RM, Cantwell MC, Pellletier MC, Ho KT, Serbst JR, Cook HF, and Kuhn A. 2000. Development of a Toxicity Identification Evaluation Procedure for Characterizing Metal Toxicity in Marine Sediments, Environmental Toxicology and Chemistry, 19(4): 982-991.

Burton GA Jr. (ed.). 1992. Sediment Toxicity Assessment, Lewis Publishers Inc. Chelsea, MI, 457 p.

Burton GA. 1998. Assessing aquatic ecosystems using porewaters and sediment chemistry, Report prepared for AETE program, CANMET, Natural Resources Canada, Ottawa, Ontario.

Cairns MA, Nebeker AV, Gakstatter JH, and Griffis WL. 1984. Toxicity of Copper-spiked Sediments to Freshwater Invertebrates, Environ. Toxicol. Chem., 3: 435-445.

Carignan R. 1984. Interstitial Water Sampling by Dialysis: Methodological Notes, Limnol. Oceanogr., 29: 667-670.

Carignan R. and Lean DRS. 1991. Regeneration of Dissolved Substances in a Seasonally

Anoxic Lake: the Relative Importance of Processes Occurring in the Water Column and in the Sediments, Limnol. Oceanogr., 36(43): 683-707.

Carignan R. and Tessier A. 1985. Zinc Deposition in Acid Lakes: The Role of Diffusion, Science, 228: 1524-1528.

Carignan R, Rapin F, and Tessier A. 1985. Sediment Pore Water Sampling for Metal Analysis: A Comparison of Techniques, Geochim. et Cosmochim. Acta, 49: 2493-2497.

Carignan R, St-Pierre S, and Gächter R. 1994. Use of Diffusion Samplers in Oligotrophic Lake Sediments: Effect of Free Oxygen in Sampler Material, Limnol. Oceanogr., 39(2): 468-474.

Carr RS, and Chapman DC. 1995. Comparison of Methods for Conducting Marine and Estuarine Sediment Pore Water Toxicity Tests-extraction, Storage and Handling Techniques, Arch. Environ. Contam. Toxicol., 28: 69-77.

Carr RS, and Chapman DC. 1992. Comparison of Solid-phase and Pore-water Approaches for Assessing the Quality of Marine and Estuarine Sediments, Chem. Ecol., 7: 19-30.

CCME. 1994. A framework for ecological risk assessment at contaminated sites in Canada: Review and Recommendations. Canadian Council of Ministers of the Environment (CCME), Environment Canada Scientific Series N. 199. Ottawa.

Chapman PM. 1992. Sediment Quality Triad Approach. U.S. EPA Sediment Classification Methods Compendium. Office of Science and Technology, Washington, DC. 10-1-10-18.

Deaver E. and Rodgers JH Jr. 1996. Measuring bioavailable copper using anodic stripping voltammetry. Environ. Toxicol. Chem. 15: 1925-1996.

de Groot AJ and Zschuppe KH. 1981. Contribution to the Standardization of the Methods of Analysis for Heavy Metals in Sediments, Rapp. P.-v. Reun. cons. Int. Explor. Mer., 181: 111-122.

Di Toro DM. 1989. A Review of the Data Supporting the Equilibrium Partitioning Approach to Establishing Sediment Quality Criteria, report to the National Research Council (as cited in Baudo 1990).

Ecological Services for Planning (ESP). 1996. Aquatic Effects Technology Evaluation, 1996 field evaluation. Final survey report for Dome Mine, Ontario. Prepared for Aquatic Effects Technology Evaluation Program, Natural Resources Canada, Ottawa, ON.

Environment Canada. 1983. Sampling for Water Quality, Water Quality Branch, Inland

Waters Directorate, EC, Ottawa, Ontario, 55 p.

Environment Canada. 1989. Bottle Washing Procedures, National Water Quality Laboratory, Inland Water Directorate, EC, National Water Research Institute, Burlington, Ontario.

Environment Canada. 1990. Biological Test Method: Acute Lethality Test Using *Daphnia* spp., Environmental Protection Series Report, EPS 1/RM/11, Ottawa, Ontario, 57 p.

Environment Canada. 1992. Biological Test Method: Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia*, Environmental Protection Series Report, EPS1/RM/21, Ottawa, Ontario, 72 p.

Environment Canada. 1992c. Biological Test Method: Fertilization Assay Using Echinoids (Sea Urchins and Sand Dollars), Environmental Protection Series Report, EPS 1/RM/27, Ottawa, Ontario, 97 p.

Environment Canada. 1992. Biological Test Method: Growth Inhibition Test Using the Freshwater Alga *Selenastrum capricornutum*, Environmental Protection Series Report, EPS 1/RM/25, Ottawa, Ontario, 41 p.

Environment Canada. 1994. Guidance Document on Collection and Preparation of Sediments for Physiochemical Characterization and Biological Testing, Environmental Protection Series Report EPS 1/RM/29. Ottawa.

Environment Canada. 1997. Test for Growth and Survival in Sediment Using the Freshwater Amphipod *Hyalella azteca*, Environmental Protection Series Report, EPS 1/RM/33, Ottawa, Ontario, 123 p.

Environment Canada. 1997. Test for Growth and Survival in Sediment Using Larvae of Freshwater Midges (*Chironomus tentans, Chironomus riparius*), Environmental Protection Series Report, EPS 1/RM/32, Ottawa, Ontario, 131 p.

Environment Canada. 1997. Recommended procedures for adjusting salinity of effluent samples for marine sublethal toxicity testing conducted under Environmental Effects Monitoring (EEM) programs.

Edmunds WM, and Bath AH. 1976. Centrifuge Extraction and Chemical Analysis of Interstitial Waters, Environ. Sci. Tech., 10(5): 467-492.

Emerson S, Jahnke R, Bender M, Froelich P, Klinkhammer G, Bowser C, and Setlock G. 1980. Early Diagenesis in Sediments from the Eastern Equatorial Pacific, I. Pore Water Nutrient and Carbonate Results, Earth Planet. Sci. Lett., 49:57 (as cited in Adams, 1991).

ESG International Inc. (ESG). 1999. AETE Synthesis Report of Selected Technologies

for Cost-Effective Environmental Monitoring of Mine Effluent Impacts in Canada, AETE Project 4.1.4, March 1999, CANMET, Natural Resources Canada, Ottawa, Ontario.
Froelich PN, Klinkhammer GP, Bender ML, Luedtke NA, Heath GR, Cullen D, Dauphin P, Hammond D, Hartmann B, and Maynard V. 1979. Early Oxidation of Organic Matter in Pelagic Sediments of the Eastern Equatorial Atlantic: Suboxic Diagenesis, Geochim. Cosmochim. Acta, 43: 1075-1090.

Giesy JP and Hoke RA. 1990. Freshwater Sediment Quality Criteria: Toxicity Bioassessment, pp. 265-348, in: Sediments: Chemistry and Toxicity of In-place Pollutants, Baudo R, Giesy JP and Muntau H (eds.), Lewis Publishers, Inc. Chelsea, MI, 405 p.

Glass GE and Poldoski JE. 1975. Interstitial Water Components and Exchange Across the Water Sediment Interface in Lake Superior, Berh. Int. Verhin. Limnol. 19: 405-420.

Golterman HL, Sly PG and Thomas RL. 1983. Study of the Relationship Between Water Quality and Sediment Transport, published by the United Nations Educational, Scientific and Cultural Organization (UNESCO), Mayenne, France, 231 p.

Green RH. 1989. Power Analysis and Practical Strategies for Environmental Monitoring, Environ. Res., 50: 195-205.

Håkanson L and Jansson M. 1983. Principles of Lake Sedimentology, Springer-Verlag, Berlin, GD, 316 p.

Hargrave BT, Phillips GA, Doucette LI, White MJ, Milligan TG, Wildish DJ and Cranston RE. 1995. Biogeochemical observations to assess benthic impacts of organic enrichment from marine aquaculture in the Western Isles region of the Bay of Fundy, Can. Tech., Rep. Fish. Aquat. Sci. No. 2062. 159 pp.

Harkey GA, Landrum PF, and Kaine SJ. 1994. Comparison of Whole-sediment, Elutriate and Pore-water Exposures for Use in Assessing Sediment-associated Organic Contaminants in Bioassays, Environ. Toxicol. Chem., 13(8): 1315-1329.

Hertkorn-Obst U, Wendeler H, Feuerstein T, and Schmitz W. 1982. A Device for Sampling Interstitial Water Out of River and Lake Beds, Environ. Technol. Lett., 3:263.

Hesslein RH. 1976. An *in-situ* Sampler for Close Interval Pore Water Studies, Limnol. Oceangr., 21: 912-914.

Hines ME, Knollmeyer SL and Tugel JB. 1989. Sulfate Reduction and Other Sedimentary Biogeochemistry in a Northern New England Salt Marsh, Limnol. Oceanogr., 34: 578-584.

Holland PT, Hickey CW, Roper DS and Trower TM. 1993. Variability of Organic Contaminants in Inter-tidal Sandflat Sediments from Manukau Harbour, New Zealand,

Arch. Environ. Contam. Toxicol., 25: 456-463.
Howes BL, Dacey JWH and Wakeham SG. 1985. Effects of Sampling Technique on Measurement of Pore Water Constituents in Salt Marsh Sediments, Limnol. Oceanogr., 30(1): 221-227.

Jahnke RA. 1988. A Simple, Reliable and Inexpensive Pore-water Sampler, Limnol. Oceanogr., 33:483-487.

Jenne EA and Zachara JM. 1987. Factors Influencing the Sorption of Metals, pp. 83-98, in: Fate and Effects of Sediment-bound Chemicals in Aquatic Systems: Proceedings of the 6th Pellsonton Workshop, Florissant, CO, Aug. 12-17, 1984, Pergamon Press, New York, NY.

Keith LH. 1992. Environmental Sampling and Analysis: A Practical Guide, Lewis Publishers, Inc., Chelsea, MI, 143 p.

Kelly CA, Rudd JWM, Furutani A and Schindler DW. 1984. Effects of Lake Acidification on Rates of Organic Matter Decomposition in Sediments, Limnol. Oceanogr., 29: 687-694.

Klinkhammer GP. 1980. Early Diagenesis in Sediments from the Equatorial Pacific. II. Pore Water Metal Results, Earth Planet. Sci. Lett., 49: 81-101.

Knezovich JP and Harrison FL. 1988. The Bioavailability of Sediment-sorbed Chlorobenzenes to Larvae of the Midge, *Chironomous decorus,* Ecotoxicol. Environ. Safety, 15: 226-241.

Knezovich JP and Harrison FL. 1987. The Bioavailability of Sediment-sorbed Organic Chemicals: A Review, Water Air Soil Pollut., 32: 233-245.

Kriukov PA and Manheim FT. 1982. Extraction and Investigative Techniques for Study of Interstitial Waters of Unconsolidated Sediments: A Review, pp. 3-26, in: The Dynamic Environment of the Ocean Floor, Fanning KA and Manheim FT (eds.), Lexington Books, Washington, DC.

Lee GF and Jones RA. 1984. Water Quality Significance of Contaminants Associated with Sediments: An Overview, pp. 1-34, in: Fate and Effects of Sediment-bound Chemicals in Aquatic Systems, Dickson KL, Maki AW and Brungs WA (eds.), Pergamon Press, New York, NY.

Levi I and Novicki TW. 1972. Interfering GLC Peaks from Materials and Chemicals in Pesticide Residue Analysis, Bull. Environ. Contam. Toxicol., 7: 193-199.

Lyons WB, Gaudette HE and Smith GM. 1979. Pore Water Sampling in Anoxic Carbonate Sediments: Oxidation Artifacts, Nature, 277: 48-49.

MacLean RS, Borgmann U, and Dixon DG. 1996. Bioaccumulation kinetics and toxicity of lead in *Hyalella azteca* (Crustacea, Amphipoda). Can. J. Fish. Aquat. Sci. 53: 2210-2220.

Malcolm SJ, Kershaw PJ, Lovett MB and Harvey BR. 1990. The Interstitial Water Chemistry of ^{239,240}Pu and ^{241}Am in the Sediments of the North-East Irish Sea, Geochim. Cosmochim. Acta., 54: 29-35.

Manheim FT. 1976. Interstitial Waters of Marine Sediment, p. 115, in: Chem. Oceanogr., Volume 6, Riley JP and Chester R (eds.). Academic Press, New York, NY.

Martens CS and Klump JV. 1980. Biochemical Cycling in an Organic-rich Coastal Marine Basin. I. Methane Sediment-water Exchange Processes, Geochim. Cosmochim. Acta., 44: 471-490.

Mayer LM. 1976. Chemical Water Sampling in Lakes and Sediments with Dialysis Bags, Limnol. Oceanogr., 21: 909-911.

MOEE. 1993. Guidelines for the protection and management of aquatic sediment quality in Ontario. Ministry of the Environment and Energy. Toronto. 27 pp.

MOEE. 1997. Sediment and biological assessment of Canagaigue Creek at the Uniroyal Chemical Ltd. Plant, Elmira, Ontario. 1995-1996. Ministry of Environment and Engergy. 37 pp + appendices.

Montgomery JR, Price MT, Holt J and Zimmerman C. 1981. A Close-interval Sampler for Collection of Sediment Pore Waters for Nutrient Analysis, Estuaries, 4: 75-77.

Montgomery JR, Zimmerman CF and Price MT. 1979. The Collection, Analysis and Variation of Nutrients in Estuarine Pore Water, Estuarine Coastal Mar. Sci., 9: 203-214.

Mount DR, Gulley DD, Russell J, Garrison TD and Evans JM. 1997. Statistical models to predict the toxicity of major ions to *Ceriodaphnia dubia*, *Daphnia magna* and *Pimephales promelas* (fathead minnows). Environ. Toxicol. Chem. 16: 2009-2019.

Mudie PJ, Piper DJW, Rideout K, Robertson KR, Schafer CT, Vilks G and Hardy IA. 1984. Standard Methods for Collecting, Describing, and Sampling Quaternary Sediments at the Atlantic Geoscience Centre, Geological Survey of Canada, Bedford Institute of Oceanography, Dartmouth, NS, 38+ p. (1984).

Mudroch A and MacKnight SD. 1991. CRC, Handbook of Techniques for Aquatic Sediments Sampling, CRC Press, Boca Raton, FL, 210 p.

Murray JW and Grundmanis V. 1980. Oxygen Consumption in Pelagic Marine Sediments, Science, 209: 1527-1530 (1980).

Nath BN, Rajaraman VS and Mudholkar AV. 1988. Modified Interstitial Water Squeezer for Trace Metal Analysis, Indian J. Marine Sci., 17: 71-72.

Nebeker AV, Cairns MA, Gakstatter JH, Malueg KW, Schuytema GS and Krawczyk DR. 1984. Biological Methods for Determining Toxicity of Contaminated Freshwater Sediments to Invertebrates, Environ. Toxicol. Chem., 3: 617-630.

Neumann PTM, Borgmann U and Norwood WP. 1999. The effect of gut clearance on metal body concentrations in *Hyalella azteca*. Environ. Toxicol. Chem. 18: 976-984.

Novicki HG, Kieda CA, Devine RF, Current V and Schaefors TH. 1979. Identification of Organic Compounds Extracted from Paper and Glass Soxhlet Thimbles, Anal. Lett., A12: 769-776.

Patrick WH Jr, Gambrell RP and Khalid RA. 1977. Physicochemical Factors Regulating Solubility and Bioavailability of Toxic Heavy Metals in Contaminated Dredged Sediment, J. Environ. Sci. Health., A12: 475-492.

Plumb RH Jr. 1979. Sampling, Preservation, and Analysis of Sediment Samples: State-of-the-Art Limitation, presented at The Fifth US/Japan Experts Meeting in The Management of Sediments Containing Toxic Substances, New Orleans, LA (November 26-28, 1979).

Reeburgh WS. 1967. An Improved Interstitial Water Sampler, Limnol. Oceanogr. 12: 163-165.

Rutledge PA and Fleeger J. 1988. Laboratory Studies on Core Sampling with Application to Subtidal Meiobenthos Collection, Limnol. Oceanogr., 33: 274-280.

Saager PM, Sweerts JP and Ellermeijer HJ. 1990. A Simple Pore-water Sampler for Coarse, Sandy Sediments of Low Porosity, Limnol. Oceanogr., 35(3): 747-751.

Sayles FL and Manheim FT. 1975. Interstitial Solutions and Diagenesis in Deeply Buried Marine Sediments: Results from the Deep Sea Drilling Project, Geochim. Cosmochim. Acta., 39: 103-127.

Sayles FL, Wilson TRS, Hume DN and Mangelsdorf PC Jr. 1973. In-situ Sampler for Marine Sedimentary Pore Waters: Evidence for Potassium Depletion and Calcium Enrichment, Science, 181: 154-156.

Schuytema GS, Nelson PO, Malueg KW, Nebeker AV, Krawczyk DF, Ratcliff AK and Gakstatter JH. 1984. Toxicity of Cadmium in Water and Sediment to *Daphnia magna*. Environ. Toxicol. Chem., 3: 293-308.

Shaner SW and Knight AW. 1985. The Role of Alkalinity in the Mortality of *Daphnia magna* in Bioassays of Sediment-bound Copper, Comp. Biochem. Physiol., 82C: 273-277.

Simon NS, Kennedy MM and Massoni CS. 1985. Evaluation and Use of a Diffusion-controlled Sampler for Determining Chemical and Dissolved Oxygen Gradients at the Sediment-water Interface, Hydrobiologia, 126: 135-141.

Sly PG and Christie WJ. 1992. Factors Influencing Densities and Distributions of *Pontoporeia hoyi* in Lake Ontario, Hydrobiologia, 235/236: 321-352.

Tessier A, Carignan R, Dubreuil B and Rapin F. 1989. Partitioning of Zinc Between the Water Column and the Oxic Sediments in Lakes, Geochim. Cosmochim. Acta., 53: 1511-1522.

U.S. Environmental Protection Agency (U.S. EPA). 1989. Generalized methodology for conducting industrial toxicity reduction evaluations. EPA-600/2-88/070.

U.S. Environmental Protection Agency (U.S. EPA). 1991a. Methods for aquatic toxicity identification evaluations: Phase I toxicity characterization procedures. EPA-600/6-91/003.

U.S. Environmental Protection Agency (U.S. EPA). Sediment toxicity identification evaluation: 1991b. Phase I (Characterization), Phase II (Identification) and Phase III (Confirmation) modifications of effluent procedures. EPA-600/6-91/007.

U.S. Environmental Protection Agency (U.S. EPA). 1991c. Toxicity identification evaluation: characterization of chronically toxic effluents, Phase I. EPA-600/6-91/005.

U.S. Environmental Protection Agency (U.S. EPA). 1993a. Methods for aquatic toxicity identification evaluations: Phase II toxicity identification procedures for samples exhibiting acute and chronic toxicity. EPA-600/R-92/080.

U.S. Environmental Protection Agency (U.S. EPA). 1993b. Methods for aquatic toxicity identification evaluations: Phase III toxicity confirmation procedures for samples exhibiting acute and chronic toxicity. EPA-600/R-92/081.

U.S. Environmental Protection Agency (U.S. EPA). 1986. Test Methods for Evaluating Solid Waste-Physical and Chemical Methods, USA Environemntal Protection Agency, 1986, Method 9060.

van de Guchte C and Maas-Diepeveen JL. 1988. Screening Sediments for Toxicity: A Water-concentration Related Problem, in: Proceedings of the 14th Annual Aquatic Toxicity Workshop, November 1-4, 1987, Toronto, Ontario, Canada, Can. Tech. Rep. Fish. Aquat. Sci., 1607: 81-91.

van der Loeff RMM. 1980. Time Variation in Interstitial Nutrient Concentrations at an Exposed Subtidal Station in the Dutch Wadden Sea, Neth. J. Sea Res., 14: 123-143.

van Raaphorst W. and Brinkman AG. 1985. The Calculation of Transport Coefficients of Phosphate and Calcium Fluxes Across the Sediment-water Interface, from Experiments with Undisturbed Sediment Cores, Water Sci. Technol., 17: 941-951.

Viel M, Barbanti A, Langone L, Buffoni G, Paltrinieri D and Rosso G. 1991. Nutrient Profiles in the Pore Water of Deltaic Lagoon: Methodological Considerations and Evaluation of Benthic Fluxes, Estuar. Coastal Shelf Sci., 33: 361-382.

Warren LA, Tessier A and Hare L. 1998. Modelling cadmium accumulation by benthic invertebrates in situ: The relative contributions of sediment and overlying water reservoirs to organism cadmium concentrations. Limnol. Oceanogr. 43: 1442-1454.

Warwick RM and Clarke KR. 1991. A Comparison of Some Methods for Analysing Changes in Benthic Community. J. Mar. Biol. Ass. U.K. 71: 225-244.

Wentworth CK. 1922. A scale of grade and class terms for clastic sediments; *Journal of Geology*, 30: 377-392.

Whiticar MJ. 1982. Determination of Interstitial Gases and Fluids in Sediment Collected with an in-situ Sampler, Anal. Chem., 54: 1796-1798.

Word JQ, Ward JA, Franklin LM, Cullinan VI and Kiesser SL. 1987. Evaluation of the Equilibrium Partitioning Theory for Estimating the Toxicity of the Non-polar Organic Compound DDT to the Sediment Dwelling Amphipod *Rhepoxynius abronius*, Final report submitted by Battell Washington Environmental Program Office to U.S. Environmental Protection Agency, Criteria and Standards Division, 60 p.

8.6 Appendix 8.1: Details on In-situ Collection of Pore Water

In shallow waters (e.g., <10 m deep), pore water has been collected by divers directly from fine-grained sediments with long-needled syringes (Bauer et al. 1988) or a tube, either by hydrostatic pressure or vacuum, after passage through a filter (Brinkman et al. 1982; Buddensiek et al. 1990). Only relatively small volumes (5 to 10 mL) can be collected at one time from a given depth. In shallow estuarine waters (0.5 m deep), a water sampler was developed to repeatedly collect anoxic pore water from estuarine sediments (Montgomery et al. 1979; 1981). Sample volumes between 50 and 75 mL were collected by vacuum in an inert atmosphere (provided by an inert gas which periodically displaced the collected sample). A similar system was designed for use in rivers and lakes (Hertkorn-Obst et al. 1982). In deeper marine waters (20 to 1800 m), a stainless-steel tube sampler which functioned like a giant syringe was developed to collect pore water simultaneously at several depths (Sayles and Manheim 1975). This design was modified to collect pore water from shallow, subtidal marine environments (van der Loeff 1980). Howes et al. (1985) collected pore water in shallow marshes from different sediment depths by modifying the design of Montgomery et al. (1979) but only 5 to 10 mL could be sampled at one time from the depth profile.

All of these methods for collecting pore water rely on water suction or displacement with inert gases, and filtering. Most other methods rely on diffusion principles. The dialysis samplers or "peepers" have membranes that differ with respect to their composition and pore size, which confer some solute selectivity (e.g., 0.2 µm polysulfone or polycarbonate membranes, 0.45 µm PVC membranes, 3.0 µm porous Teflon) (Adams 1991). Cellulose-based dialysis membranes are susceptible to microbial attack and should not be used (Martens and Klump 1980). Degassing of PVC membranes is desireable before use in peepers (Carignan et al. 1994). The dialysis sampler is usually deoxygenated in distilled, or demineralized water that has been purged with nitrogen (Hesslein 1976; Kelly et al. 1984; Carignan and Tessier 1985; Tessier et al. 1989; Viel et al. 1991) or in a nitrogen-purged solution of sodium chloride (Simon et al. 1985). Degassing with nitrogen for 24 to 48 h in an equilibration chamber is generally used to deoxygenate the dialysis sampler. The sampler should be placed directly from the equilibration chamber into the sediment (usually by a diver), where it is allowed to equilibrate for a period of time (3 to 20 days) before the pore water is collected from the recovered dialysis sampler (i.e., either by syringe or attached collecting tubes). The equilibration time should be experimentally determined, *a priori*, by daily sampling and analyses of the pore water (Carignan 1984).

Experimental artifacts can result from contamination of the pore water by the sampling equipment, incomplete equilibration between the sediment pore water and the water in the dialyzer, membrane disruption, as well as interference with diffusion of ions from the development of an electrical potential across the membrane or oxidation of the pore water during collection from the dialyzer and transfer to the sample container (Carignan et al. 1985). These artifacts can be minimize by:

a) using only clean chemically inert material in the construction of the dialyzer (i.e., no rubber, neoprene, PVC, polystyrene, metallic or galvanized metallic materials);
b) constructing the dialyzers carefully with intact, sealed porous Teflon®, polycarbonate membranes, or biologically inert polysulfone;
c) using only a purified oxygen-purged equilibrator in the chamber;
d) ensuring that the equilibration time is adequate;
e) using non-metallic syringes for extraction of pore water;
f) using clean, pretreated storage or sample containers of the appropriate material (Section 2.8); and
g) transferring the pore water from the dialyzer to the sample container in an oxygen-free atmosphere (i.e., glove box with argon or nitrogen gas).

Viel et al. (1991) compared the iron and nutrient ammonia, phosphate, silica (NH_3, $P0_4$, $Si0_2$) concentrations in pore water, collected by centrifugation at 500 rpm for 20 min, with the concentrations in pore water collected *in situ* by dialyzers. They found that the concentrations of $P0_4$ and Si were comparable in the two types of water, but that concentrations of iron and ammonia differed, probably because of manipulation of the core samples during transfer to the centrifugation tubes, and/or spatial heterogeneity.

Knezovich and Harrison (1987) compared the concentrations of chlorobenzene in pore water collected *in situ* from the surficial layers of sediment (e.g., 0 to 8 mm), by a modified syringe, with that found in pore water collected by compression methods and centrifugation. The loss of the volatile compound was greatly reduced by the *in situ* extraction method.

Dialyzers have the advantage of collecting pore water samples free of temperature, pressure and oxidation artifacts, but require a relatively-long equilibration period. The suction samplers can recover samples that are also relatively free from the same artifacts, but are more time consuming to operate, more expensive to construct, and have limited depth resolution. Since the two methods of sampling pore water in situ (e.g., water suction versus dialyzers) have only recently been directly compared, one method cannot be recommended over another.

8.7 Appendix 8.2: Details on Recovery of Pore Water from Bulk or Core Sediment Samples

The materials from which the components of the squeezing device are made can affect the chemistry of the pore water. Devices made of stainless steel are suitable for collecting pore water for the determination of anions, alkali metals and alkaline earths, but not for iron and other transition elements. These elements are subject to contamination from stainless steel (Adams 1991).

Using a mechanical squeezer, Manheim (1976) observed that squeezing pressure did not significantly affect the composition of the extracted pore water. A problem that may be associated with squeezing is the formation of an impermeable cake at the bottom of the squeezing device, preventing the passage of pore water. Stepwise increases in pressure over a 30-minute period can minimize this problem (Reeburgh 1967). Nath et al. (1988) modified the Manheim squeezer to reduce the collection time from between 60 and 90 min to collect 20 mL of pore water to less than 20 min. The new design included a Teflon® liner, which prevented exposure of the sediment to the atmosphere, thus limiting oxidation effects. Jahnke (1988) also developed a simple, reliable and inexpensive whole-core squeezer to collect the pore water in gas-tight syringes connected to the core tube to minimize oxidation effects.

Klinkhammer (1980) observed high concentrations of trace metals, particularly iron, in pore water, and suggested that iron-rich, colloidal-sized particles were not retained by the filters in squeezing devices. Hines et al. (1989) found that sulphide concentrations were lower in squeezed samples when compared to samples of pore water collected *in situ*. This occurred even though a strict oxygen-deficient atmosphere was maintained during squeezing. Malcolm et al. (1990) found that concentrations of reduced plutonium (a transuranium nuclide) increased on exposure to air and they recommend that sampling and extraction of interstitial waters from anoxic sediment be conducted in a glove box to minimize compositional changes.

Chemical fractionation of a pore water sample can also occur during squeezing. Significantly high ammonium concentrations were reported in pore water in the first three aliquots (Emerson et al. 1980). Froelich et al. (1979) compared two types of squeezers and found that, for both squeezers, when calcium carbonate ($CaC0_3$)-rich sediments were squeezed through filters that had been acid-washed, pore water carbon dioxide ($C0_2$) levels were higher in the first aliquot obtained than in subsequent aliquots.

It is accepted that some pore water constituents are affected by the collection method. Dissolved organic carbon and dimethylsulphide are affected, while the effects on ammonia and sulphide (given that oxidation is prevented) are equivocal (Burton 1992). Bender et al. (1987) concluded that the composition of pore water collected by a whole-core squeezing device was very susceptible to rapid alteration by solid-solution reactions, and that the method was unsuitable for determining pore water profiles of trace metals and other particle-reactive chemicals.

Centrifugation has been compared to an *in-situ* dialysis method and was found to give similar results. The general procedure for centrifuging a sediment core sample is as follows. A horizon of a core sediment is cut, or displaced, in an oxygen-free environment (e.g., glove box, glove bag, etc.). The sediment is loaded into acid-washed polycarbonate centrifuge tubes. The tubes are tightly capped and centrifuged at a high speed while under refrigeration. The supernatant is then siphoned off (in the oxygen-free environment) and only filtered if necessary (i.e., as stipulated in the test method). The supernatant or filtrate is then stored in a polystyrene vial that has been previously acidified with hydrogen chloride (HCl) or nitric acid ($HN0_3$).

The variable aspects of the centrifugation process are the centrifuge speed and the composition and pore size of the filter used. Adams et al. (1980) examined the effect of centrifuge speeds (7,000 to 19,000 rpm) on pore water composition. Little change was found in concentrations of calcium (Ca), iron (Fe), manganese (Mn) and zinc (Zn), but ($P0_4$) concentrations doubled. Another concern is that the centrifuge speed may not be high enough to remove dispersible clays.

Processes such as adsorption of potential toxicants in the pore water to the filters and desorption of potential toxicants from the filter require that filters be pretreated and that filter blanks be used to assess the extent to which these processes might occur (Levi and Novicki 1972; Novicki et al. 1979). Filtration with glass fibre or plastic filters has been shown to remove non-polar organics. Certain dissolved hydrophobic contaminants might be adsorbed to glass fibre filters (Word et al. 1987). Other studies recommend the use of polycarbonate filters (Knezovich and Harrison 1987). Carr and Chapman (1995) compared the toxicity of pore water collected from subsamples of contaminated sediment with a Teflon® extractor and centrifugation. Five types of filters (e.g., fluorocarbon, nylon, polyester, polycarbonate, glass fibre) were used and testing demonstrated that pretreatment which involved soaking the filters in deionized water for 24 h did not adequate remove likable toxic compounds from the nylon filters in 24 h and that the glass fibre filters could not withstand the soaking. The problem with the nylon filters was overcome by simply extending the soaking period to 48 h. Different types of filters used in squeezing devices adsorbed soluble contaminants to varying degrees. Recent studies indicate that filtering pore water will substantially reduce toxicity regardless of the method of pore water extraction (e.g., centrifugation, pressure extraction, syringe extraction or dialysis) (Ankley et al. 1991). Saager et al. (1990) used a centrifuge tube specially designed for collection of pore water.

Carignan et al. (1985) compared pore water recovered from field-collected sediments by centrifugation and filtration with that collected from sediment *in situ* by dialysis. Centrifugation was conducted at two speeds (5,000 and 11,000 rpm), followed by filtration with pore sizes of 0.45, 0.2 and 0.03 µm. The concentration of calcium (Ca), magnesium (Mg), iron (Fe), manganese (Mn), chromium (Cr), cobalt (Co), nickel (Ni), copper (Cu), zinc (Zn), cadmium (Cd) and organic carbon were measured and compared. Centrifuging at 5,000 rpm and filtering through a pre-washed 0.45 µm filter resulted in pore water levels of Co, Ni, Cr, Fe and Mn comparable to those in the pore water obtained by the dialysis method. However, Cu, Zn and organic carbon levels were significantly higher in pore water obtained by the centrifuge method. Centrifuging at 11,000 rpm and filtering through 0.2 or 0.03 µm filters produced pore water levels of Cd, Co, Cr, Ni and organic carbon equivalent to those observed in the dialysis pore water. It would appear that centrifuging compares well with the *in-situ* dialysis method, at least for the elements examined in this study.

The effects of five pore water extraction methods on sediment toxicity were evaluated using sediments that were heavily contaminated with metals, and/or oil and grease, PAHs and NH_3 (Ankley et al. 1991). The extraction methods included centrifugation at low and high speeds (e.g., 250 x G and 10,000 x G, respectively) at 4°C, pressure extraction, syringe extraction, as well as dialysis. The pore water samples were collected with and without filtration. The comparative results indicate that centrifugation speed had little effect on sediment toxicity. However, it was concluded that the higher speed was more appropriate because it isolated only the bioavailable metal fraction whereas the lower speed resulted in the presence of unavailable metals in pore water. Pore water collected by centrifugation was generally as toxic or more toxic than pore water collected by the other methods. Filtering pore water decreased toxicity by between 40 and 85% (Ankley et al. 1991).

The effects of three pore water extraction methods (e.g., centrifugation, vacuum suction, and pressurized squeezing) on toxicity of contaminated sediments were evaluated (Carr and Chapman 1995). They concluded that: the toxicity of pore water samples is greatly influenced by the type of collection method; centrifugation should be used when the primary contaminants of concern are highly hydrophobic organic compounds to maximize the sensitivity of egg fertilization test; and, regardless of the method used for the initial extraction, the sample should be (re-) centrifuged prior to testing or freezing to remove suspended particulate which compromise test results.

Chemical fractionation of the pore water sample is a concern for the centrifuge method as well. Edmunds and Bath (1976) studied the potential fractionation during centrifuging of five cations (Na^+, K^+, Ca^{2+}, Sr^{2+}, Mg^{2+}) from the pore water of Cretaceous rock. Cation concentrations were measured at four times during pore water removal. Sodium and potassium showed a progressive depletion as a greater portion of the pore water was extracted, with concentrations levelling off during the extraction of the final 10 to 20% of the sample. Concentration of Ca^{2+} decreased until extraction was complete. No apparent fractionation of Sr^{2+} or Mg^{2+} was observed.

Other methods for extraction of pore water from samples of sediments include dessication procedures (Adams 1991), vacuum filtration (Glass and Poldoski 1975), and displacement techniques (Adams 1991). The displacement technique involves the use of an immiscible liquid to displace the pore water. Extraction is made with a filter press after the liquid is poured over the sediment and then pressurized (Adams 1991). Vacuum filtration with a Buchner funnel is used to extract pore water (Glass and Poldoski 1975). This is a tedious process, especially with fine-grained sediments where evaporation of the pore water becomes a major problem (Adams 1991). The potential errors involved with this technique are the solubilization of solids and the changes that occur in the sediment-water equilibria.

van Raaphorst and Brinkman (1985) used polyethylene tubing containing cotton threads to extract pore water from a sample of cored sediment. Under vacuum, 5 mL of pore water was collected over two to three days. Phosphate in the pore water did not interact with the cotton. This procedure would not be appropriate for samples where analyses for organic contaminants are desired.

Table of Contents

9. Data Assessment and Interpretation ... 9-1

9.1 Understanding the objective of the EEM program, the definition of effect, and meaning of data interpretation within EEM ... 9-1

9.2 Data Analyses ... 9-3
9.2.1 Effects on Fish ... 9-3
9.2.2 Effects on Usability of Fisheries Resources ... 9-9
9.2.2.1 Mercury in Fish Tissue ... 9-10
9.2.2.2 Tainting ... 9-11
9.2.2.3 Physical Abnormalities ... 9-12
9.2.3 Effects on Benthic Invertebrate Community ... 9-13

9.3 The Role of Power Analysis, α, β, and Effect Size in Determining Effects ... 9-20
9.3.1 Setting αand β ... 9-20
9.3.2 Power Analysis: Determination of Required Sample Size, Power, and Appropriate Effect Size ... 9-23

9.4 Alternative Methodologies for the Fish and Benthic Invertebrate Community Surveys ... 9-26

9.5 Sublethal Toxicity ... 9-28

9.6 References ... 9-29

Appendix 1.1 Summary Questions for Fish Survey ... 9-31

Appendix 1.2 Summary Spreadsheet for Fish Survey ... 9-32

Appendix 2.1 Summary Questions for Benthic Invertebrate Community Survey ... 9-34

Appendix 2.2 Summary Spreadsheet for Benthic Invertebrate Community Survey 9-34

List of Tables

Table 9-1: Effects on fish ... 9-4
Table 9-2: Endpoints (supporting response variables) to be used for supporting analyses.9-4
Table 9-3: Effects on Usability of Fisheries Resources ... 9-10
Table 9-4: Endpoints to be used for supporting analyses. ... 9-10
Table 9-5: Effects on Benthic Invertebrate Community ... 9-14
Table 9-6: Statistical procedure used to determine an effect for each of the six study designs. ... 9-14
Table 9-7: Sample sizes required to detect difference of ±2 SD ... 9-22

9. Data Assessment and Interpretation

As part of the Environmental Effects Monitoring (EEM) requirements under the Metal Mining Effluent Regulations (MMER), mines subject to the MMER are required to submit an interpretative report describing monitoring results. The purpose of this chapter is to provide general guidance on how to interpret EEM data - specifically:

1) which effect endpoints to use and report,
2) the statistical (or other) approach to use for each endpoint to determine the presence or absence of an effect,
3) the role of power analysis, α, β, and effect size in determining effects.

As outlined below, a statistically significant difference or gradient in any of the effect endpoints is to be considered an effect. These endpoints are divided into:

1) effects on fish,
2) effects on the use of fish by man (usability of fisheries resources),
3) effects on benthic invertebrate communities (indicator of fish habitat).

9.1 Understanding the objective of the EEM program, the definition of effect, and meaning of data interpretation within EEM

The current objective of the EEM program is *to evaluate the effects of mine effluent on the aquatic environment, related to protecting fish, fish habitat, and the use of fish by man. This information will be used to assess the adequacy of the MMER to protect fish, fish habitat, and the use of fisheries resources.* The underlining premise of this objective is that studies are designed to:

- utilize laboratory and field monitoring approaches which directly or indirectly assess the health of fish (fish survey), habitat impairment (benthic invertebrate community survey), and concerns relating to human consumption of fish (fish tissue analysis);
- discriminate between mine-related effects and other natural or anthropogenic stressors;
- generate relevant data which may be appropriately analyzed for spatial and temporal trends in a way that can be reliably interpreted.

Not all of these points will necessarily be addressed in a single EEM study. EEM has a tiered monitoring approach, where the results of the preceding EEM study will influence the study design of the next EEM study.

Two issues that need to be understood for the EEM program and especially for writing an interpretative report relate to: 1) the types of data analyses that are relevant and 2) what is meant by the term "interpretation". In order to address both these issues, it is important to define what is an effect.

Within EEM, an effect is defined generally *as a statistically significant difference in fish, fish usability, or benthic invertebrate community endpoints measured between an area* *exposed to effluent and a reference area or a statistically significant gradient in these endpoints from the exposure to reference area.* For fish tissue analysis (which is conducted to determine the usability of fisheries resources), measurements of total mercury that exceed 0.45 µg/g wet weight total mercury in the exposure area and that are statistically different from measurements of total mercury in fish tissue in the reference area defines an effect in fish tissue. In cases where it is not feasible to examine wild fish or field distribution of benthic invertebrates in areas exposed to effluent and reference areas, an alternative monitoring approach for fish or fish habitat may be used to determine if the effluent is causing an effect.

Given the above definition of effect, it is important to recognize that not all effects identified in EEM will represent damage to fish, fish habitat, or the usability of fisheries resources. However, effects as defined above do represent scientifically defensible differences or gradients that may reflect changes to the ecosystem associated with the effluent. As a result, detailed information on the effects, including the magnitude, geographic extent, and possible cause of the effect, will contribute to the understanding of the ecosystem and will be used in the management of the aquatic resources.

The "Guidance for Determining Follow-up Actions when Effects Have Been Identified in Environmental Effects Monitoring (EEM)" (Environment Canada 2001) provides guidance to regulatory agencies, regulated facilities, and EEM practitioners for determining follow-up actions when effects have been identified in EEM. It discusses the factors that need to be considered for determining follow-up actions, and explains the roles and responsibilities of government, industry, and stakeholders in this process. The document was developed by Environment Canada, in consultation with Fisheries and Oceans Canada.

Consequently, the outcome of determining the presence of an effect (e.g., a statistical difference in an effect endpoint, such as relative gonad size, between exposure fish and reference fish) is further monitoring-related studies (e.g., confirmation of effect, determination of cause, determination of magnitude, and delineation of the geographic extent) and the difference itself is *not* a sole determinant in resource management decision-making. Further, the role of the interpretative report within the EEM program is to summarize study results (including difficulties or confounding factors encountered), conduct applicable spatial (and, when sufficient data are available, temporal trend) analyses, specify any identified "effects", and make recommendations for subsequent EEM program monitoring. Data interpretation or the role of the report does *not* include determining the ecological, economic, or social significance of results (see further discussion in Section 4.2).

9.2 Data Analyses

EEM involves iterative phases of monitoring and reporting. At the present early stage of EEM, the monitoring design and statistical power are particularly well suited for detecting spatial effects. Therefore, unless sufficient historical data are present, statistical analyses and data interpretation should focus on spatial effects within each monitoring phase. More specifically, the data generated for each mine should be analyzed to determine whether there are significant differences in certain measurement endpoints between reference and exposure areas or along an exposure gradient (i.e., determination of effect), which is the focus of this chapter. In later monitoring phases, when sufficient temporal data becomes available, it will become more feasible to conduct temporal statistical analyses (i.e., to determine statistically whether any effects identified previously are lessening or worsening). Prior to formalized statistical analyses of temporal trends, however, it is useful to qualitatively compare spatial effects found in earlier monitoring studies to those found in later monitoring studies.

For EEM purposes, only specified data (the effect endpoints) generated from the fish survey, benthic invertebrate community survey, and fish usabilty studies will be used to assess the presence of effects. Other EEM requirements are only used to help interpret effects on fish and benthos (e.g., effluent characterization and water quality monitoring) or to help characterize any changes in effluent quality over time (e.g., sublethal toxicity testing). The endpoint tables in the following sections summarize the recommended data analysis procedures for the effect endpoints for each monitoring requirement (Tables 1-3). Also refer to the relevant sections of this Guidance Document for further details. Many of the data interpretation issues are the same for the fish survey, fish usability, and benthic invertebrate community sections that follow (e.g., assumptions and interpretation of statistical techniques common to more than one of these sections). These issues are discussed in more detail in Section 3.1. Reference to Section 3.1 will, therefore, aid with understanding the steps outlined in Sections 3.2 and 3.3.

9.2.1 Effects on Fish

Table 9-1 and Table 9-2 outlines the effect endpoints, supporting response variables, and appropriate statistics that are applicable for the fish survey. Differences between sexes in growth rate, body weight, condition factor, gonad size, and liver size are common due to differences in overall energetic requirements between male and female fish. Therefore, for all parameters, sexes should initially be treated separately when estimating variability. In addition, sexually immature fish should not be mixed with sexually mature fish for analyses.

Table 9-1: **Effects on fish.** Endpoints to be used for determining "effects" as designated by statistically significant differences between exposure and reference areas.[1]

Type of Response	**Effect Endpoint**	**Statistical Procedure[2] (dependent on endpoint)**
Survival	Age	ANOVA
Energy Use	Size-at-age (body weight against age)	ANCOVA
	Relative gonad size (gonad weight against body weight)	ANCOVA
Energy Storage	Condition (body weight against length)	ANCOVA
	Relative liver size (liver weight against body weight)	ANCOVA

Table 9-2: Endpoints (supporting response variables) to be used for supporting analyses.

These analyses are for informational purposes and significant differences between exposure and reference areas are not necessarily used to designate an effect.[1]

Type of Response	**Supporting Response Variable**	**Statistical Procedure[2] (dependent on endpoint)**
Energy Use	Body weight (whole)	ANOVA
	Length	ANOVA
	Size-at-age (length against age)	ANCOVA
	Relative gonad size (gonad weight against length)	ANCOVA
	Relative fecundity (# of eggs/female against body weight)	ANCOVA
	Relative fecundity (# of eggs/female against length)	ANCOVA
	Relative fecundity (# of eggs/female against age)	ANCOVA
Energy Storage	Relative liver size (liver weight against length)	ANCOVA
	Relative egg size (mean egg weight against body weight)	ANCOVA
	Relative egg size (mean egg weight against age)	ANCOVA

[1] For the ANCOVA analyses, the first term in parentheses is the endpoint (dependent variable, Y) that is analyzed for an effluent effect. The second term in parentheses is the covariate, X (age, weight, or length).

[2] ANOVA - Analysis of Variance; ANCOVA - Analysis of Covariance

Descriptive statistics should be reported in graphical and tabular format for all variables (mean, median, minimum, maximum, standard deviation, standard error and sample size). These should be reported for each area sampled (e.g., reference area and exposure area). For each of the ANOVA and ANCOVA analyses, the level of replication (sample size, n) is the number of individual fish. The minimum sample size recommended is 20 sexually mature fish per sex (and an additional 20 sexually immature fish if small-bodied fish species are being sampled) for each of the two sentinel fish species for each reference and exposure area. A power analysis should be conducted to determine sample size if the appropriate data are available.

In addition to descriptive statistics, the main analyses used to determine whether statistically significant effects have occurred should follow these three steps of data inspection, analysis and interpretation:

1) The data should be inspected to see whether they satisfy the assumptions of ANOVA or ANCOVA. These procedures are robust enough to allow for moderate violations of some assumptions and, in some cases, data transformation will help to remedy departures from the assumptions. In cases where data transformations do not sufficiently rectify departures from the assumptions, it may be necessary to use nonparametric procedures, in which case the methods of power analysis discussed in Section 4 would not apply. These issues are further discussed below, and the standard statistical texts (e.g., Sokal and Rohlf, 1995) should be consulted for a more complete discussion.
2) Following inspection of the data and any necessary transformations, the actual statistical comparisons are carried out. It should be noted that the ANCOVA is basically a two step procedure consisting of:
   a) determining whether the slopes are approximately parallel (given that they are significantly different from zero to begin with),
   b) if the slopes are parallel, going on to determine whether the elevations of the regressions are significantly different. This procedure is discussed more fully below.
3) After the statistical comparisons are made, the key results for the effect endpoints (Table 9-1) will be presented in a clear fashion so as to indicate whether there has been a statistical effect and, if so, what is the nature of the effect (including the direction and magnitude of the effect). A statistical effect is declared if the p value exceeds the *a priori* α value determined as outlined in Section 4.1.

Area differences in age, weight, and length can be tested using ANOVA (Table 9-1 and 9-2). The assumptions for ANOVA are that:

- the data for reference and exposure populations are normally distributed,
- the variances are equal between the reference and exposure populations,
- the error terms are independently distributed.

Data should be inspected for normality and homogeneity of variance before applying these parametric procedures. However, univariate normal distribution-based statistical methods are often fairly robust and can support moderate violations of the assumptions. Transformation of original data can help normalize the data or homogenize the variances.

Logarithmic transformations are often preferred because biological measures are frequently considered to operate on a logarithmic or exponential scale (Peters, 1983). If the transformations are unable to produce data that approximately meet the assumptions, then it may be necessary to use non-parametric statistics.

Lack of independence can occur when, for example, one person collects all the data from the exposure area while another person collects data from the reference area. This can bias the data if the two persons consistently use slightly different sampling or sorting protocols. Generally, these kinds of problems can only be remedied by changing the method of conducting the sampling to remove the sources of bias.

The remaining endpoints are analyzed using regressions of one variable on another using Analysis of Covariance (ANCOVA) (Table 9-1 and 9-2). ANCOVA combines the features of analysis of variance and regression, and can be used to compare regressions among treatments (i.e., control vs. impact areas). The purpose of using ANCOVA is to factor out the influence of the covariate, X, on the endpoint, Y, before determining whether there has been a significant effect of the effluent on Y for a given X. For example, the mean squared error (MSE) of the ANCOVA model can be used to provide an estimate of variance of the fish parameters, particularly at larger sample sizes (see Chapter 4).

Assumptions of ANCOVA are that:

- the residuals are normally and independently distributed with zero mean and a common variance,
- the independent variable (covariate) is fixed and measured without error,
- the relationship has the form specified (linear regression),
- the slopes of the regression lines among areas are equal.

As for ANOVA, transformations of the original data often help improve normality and homogenize variances. The assumption that the independent variable is fixed is frequently violated and Draper and Smith (1981) discuss the consequences of this violation. A non-fixed independent variable is likely to prove problematic mainly in situations where the range of the independent variable is very small - i.e., when the range in size (or age) of the fish included in the regression is very small. In this case (very narrow size or age range), however, there is little to be gained by using ANCOVA with size or age as a covariate, and the data would better be analyzed as a simple ANOVA comparison of the exposure to reference area (i.e., no need to factor out the influence of the covariate). The assumption of a linear relationship can be tested for samples with multiple observations at different values of the independent variable. This may be possible for discrete variables such as age, but not for continuous independent variables such as body weight. At a minimum, linearity should be verified by visual inspection. Linearity can often be improved by transformation (e.g., log-log). The regression plots should also be inspected to ensure that the slopes are not unduly influenced by outliers.

A key assumption of ANCOVA is that the slopes of the regression lines for the reference vs. exposure areas are equal. Therefore, the first part of an ANCOVA analysis is to test for differences in slopes between areas. A significant interaction term in the ANCOVA for covariate X vs. area (e.g., age*area or size*area) indicates significantly different slopes. In cases where the slopes are not significantly different (i.e., interaction term not significant), this indicates that the regression lines are approximately parallel to each other. Using the size-at-age ANCOVA as an example, parallel slopes would indicate that weight gain over age is similar for both areas. The next step in this example is to proceed with the ANCOVA model and test for differences in adjusted means (elevation) to investigate whether fish are proportionately heavier at any age in one area than in another.

It is possible that the slopes of regressions may differ. For example, fish from the reference area may be gaining weight more rapidly with increasing age (steeper slope) than fish from the exposure area. If the slopes of the regressions are significantly different, the ANCOVA analysis can not be completed. In this case, using the size-at-age example, the effect would not be a proportional difference in weight at any age, but rather that the rate of weight gain with increasing age is significantly different among areas. That is, an effect would be determined as a significant difference in slope among areas rather than a significant difference in elevation. For this situation, it is also a good idea to plot the separate regression lines to obtain a better qualitative understanding of the size-at-age relationship for each area over the entire data range of the X covariate (e.g., where do the lines intersect?). It should be noted that, even when the slopes of the regressions significantly differ among areas, it is still possible to make further comparisons over a particular range of values for the X covariate (i.e., a particular age or size range) (Sokal and Rohlf, 1995). This kind of comparison would be appropriate if it is judged that that particular age or size range is of particular concern.

It is also preferable that the range of the independent variable should be approximately the same for each "treatment" (i.e. area). This may be difficult to assure in practice, but the violation of this should be considered when interpreting results from such cases. For example, if the size range used as the X covariate for the reference area does not show much overlap with the size range for the exposure area, use of the ANCOVA results requires the assumption that the regression slopes would still be parallel for overlapping size ranges.

In addition to the above considerations, the suitability of the fish species that were used should also be evaluated, based on the site-specific nature of the results and the site-specific concerns about residency. Of specific concern is whether the fish from reference and exposure areas were mixing and, if so, whether the differences in exposure were great enough to make statistical comparisions of the collected fish meaningful. Chapter 4 provides further guidance on making these evaluations.

To summarize, the overall procedure which should be followed and reported can be divided into the following stages: 1) preparing the analyses, 2) initial summary statistics, 3) ANOVA analyses, 4) ANCOVA analyses, and 5) power analyses.

1) Preparing the analyses: this step includes data entry and QA/QC, including a check for data entry errors, and confirmation that all fish which are assumed to be adults are

undergoing gonadal development for the next spawning season. Data that will be necessary for interpretation should also be summarized, including significant confounding factors. Other significant information should include the gear used and catch per effort calculations. This step should also include an examination of the exposure data, and determination of whether there was a possibility of exposure area fish moving between areas, or whether the reference area fish were exposed.

2) Summary statistics should be calculated by species and sex, and should include means, standard deviations, minimum and maximum values and sample sizes. These should be reported in both graphical and tabular format for each area (reference and exposure area(s)) being summarized. The data should also be examined for normality and equality of variances (basic statistical assumptions).
3) ANOVA analyses should examine whether there was a statistically significant difference between exposure and reference areas in age, weight, or length for either sentinel species or sex (or juveniles, if included). The ANOVA output and interpretation should clearly outline the nature of any "effects" that have been detected (e.g., direction and magnitude), including: a) the level of significance (*p*) for the test statistic, b) percent difference between areas, and c) an indication of whether the exposure area mean(s) was larger or smaller than reference, when significant differences exist. See Section 4.0 for guidance on what level of significance (a) to compare *p* against to judge whether an effect has occurred.
4) ANCOVA analyses should be performed for size-at-age, condition factor, relative gonad size, relative fecundity, relative egg size, and relative liver size. The data should be plotted and the regressions inspected for violation of the assumptions outlined above. The initial analyses should also examine whether the regression slopes are significantly different than zero - that is, whether the endpoint varies systematically with the covariate (size or age). If the endpoint does not vary with the covariate, ANCOVA offers little advantage over ANOVA. If slopes are significantly different than zero and are equal among reference and exposure areas, the ANCOVA should proceed to analyze for a difference in elevation (adjusted mean) among areas. If the slopes are significantly different than zero, but the slopes are also significantly different among reference and exposure areas, then the ANCOVA does not proceed and the effect of exposure is a change in slope. Analogous to an ANOVA, the ANCOVA output and interpretation should outline the nature of any effects, including: a) the level of significance (*p*) for the test statistics for differences in slope or adjusted mean between areas, b) percent difference in adjusted mean between areas (if slopes were equal) or percent difference in slope (if slopes were significantly different), and c) an indication of whether the exposure area adjusted mean(s) or slopes(s) was larger or smaller than reference, when differences exist. See Section 4.0 for guidance on what α to compare *p* against.
5) The statistical power of the reference vs. exposure comparisons for the current monitoring study should be determined and compared to the power calculations done before the current study was performed, to see if expected power to detect the predetermined effect size was achieved. This is particularly important if any of the relevant parameters that could affect power (i.e., sample size, a, effect size, standard deviation) have changed since the beginning of the study. In addition, these

calculations should be used to make sample size recommendations for subsequent monitoring study. See Section 4.2 for further guidance.

9.2.2 Effects on Usability of Fisheries Resources

The purpose of examining the usability of fisheries resources is to determine whether the effluent has altered fish in such a way as to limit their use by humans. Fish usability can be affected by altered appearance, altered flavour or odour (tainting), or tissue contaminant levels which exceed consumption guidelines for human health and levels found in the reference area. Table 9-3 and Table 9-4 outlines the effect endpoints and appropriate statistics (or guideline levels) that are applicable for usability of fisheries resources.

Table 9-3: Effects on Usability of Fisheries Resources.

Endpoints to be used for determining "effects" as designated by statistically (or otherwise) significant differences between exposure and reference areas.

Effect Endpoint	**Statistical Procedure (dependent on endpoint)**
Contaminants in fish tissue (mercury if applicable)	ANOVA and evaluate against tissue guideline levels

Table 9-4: Endpoints to be used for supporting analyses.

These analyses are for informational purposes and significant differences between exposure and reference areas are not necessarily used to designate an effect.

Supporting Response Variables	**Statistical Procedure (dependent on endpoint)**
Physical abnormalities	Chi-square (separate test done for each class of abnormality; number of tests will depend on how many classes of abnormalities are present in the fish collected)
Tainting	ANOVA

9.2.2.1 Mercury in Fish Tissue

One of the methods for evaluating fish usability is by measuring concentrations of contaminants of concern in tissue from fish collected from the exposure and the reference areas. Contaminants may be identified as a concern if they are present in the effluent and there are applicable human health consumption guidelines for those contaminants. Local consumption and commercial fisheries should guide which fish species and edible tissues (e.g., liver, kidney, bones, flesh, or even entire fish) should be analyzed. Chapter 4 provides further guidance on methods for determining which (if any) contaminants should be included in the analyses. This determination depends, in part, on previously collected data on contaminant levels in fish tissue and the effluent.

Mines are required to measure levels of mercury in fish tissue if mercury is detected in the effluent (during effluent characterization – Chapter 6) above 0.10 µg/L. An effect in fish tissue, as defined in the MMER, means measurements of total mercury that exceed 0.45 µg/g wet weight total mercury in the exposure area and that are statistically different from measurements of total mercury in fish tissue in the reference area defines an effect in fish tissue. Other potential metal mine-related contaminants of concern on a site-specific basis include copper, zinc, manganese, cyanide, radium, and uranium.
Chapter 4 recommends that tissue analyses be performed on five composite samples (each composed of at least eight individual fish) of a single species (preferably one sex) for each of both areas. That is, the sample size (n) for the ANOVA is five. This would be sufficient replication to detect an effect size of ±2 SD at power = 0.9, if α and β are set at 0.1 (see Section 4.0). Thus, careful consideration should be given to the appropriate effect size to use for the particular contaminant of concern and whether increased replication may be justified. If lesser effect sizes (i.e., less than 2 SD) or greater power levels are decided to be more appropriate for the contaminant, it will be necessary to increase sample size by analyzing more composite samples.

Percent lipid and percent moisture should also be reported for each tissue sample. This is for informational purposes only to aid in data interpretation. Statistical differences in percent lipid or moisture does not constitute an effect.

9.2.2.2 Tainting

The definition of tainted fish used for EEM is based on the ISO 5492: 1992 Sensory Analysis Vocabulary and is defined as a "fish that contains an abnormal odour or flavour". The context of the definition is that the odour or flavour is introduced into the fish from outside sources and does not include any natural by-product from the deterioration in quality due to aging of fish in storage, decomposition of protein, fat, or other components, or due to microbial contamination normally found in fish.

The recommended experimental design and method of analysis is the Difference-from-Control Test. Samples of treated fish which have been exposed to effluent either under field or laboratory conditions and blind (hidden) control samples are compared with an identified control sample. The assessment of the presence of a difference based on a specified characteristic (i.e., taint) and the degree of that difference is scored using a standard sensory evaluation ballot.

Each testing session includes two reference samples and two exposure samples. Each sample will be a composite of at least three fish and preferably six or more from the same station within an area (reference or exposure), either from the same cage if cages are used, or from the same area of capture if wild caught fish are used. In other words, fish should be collected from two stations located in a reference area and two stations located in the exposure area, resulting in four groups of fish. Each of these samples is scored "blind". The effluent concentration should be approximately the same between the two exposure area stations.

At one of the reference area stations, a "known reference/control" sample is also produced. This known reference sample is to come from the same fish composite as one of the "blind reference" samples for that station. The "known reference" material is then presented with each of the "blind" samples. The flavour of the blind samples is scored in comparison to the known reference sample. Each of at least ten panelists is presented with four blind samples for scoring, in at least two sessions. They are asked to score each blind sample on a 0 to 15 scale where 0 is no difference from the known reference and 15 is extremely different.

The data are then analyzed by ANOVA, using a randomized complete block design, with panelists and sessions as blocks, and using a fixed effects model. The assumptions of ANOVA should be satisfied as outlined in Section 3.1. Data transformations should be used if necessary to satisfy the assumptions of normality and equal variances.

In summary, the minimum design includes 2 areas (reference and exposure) x 2 replicate stations per area x 10 panelists x 2 sessions. An effect has occurred if the F test for areas (reference vs. exposure) in the ANOVA is significant. Environment Canada (1998) provides further guidance on the experimental design and how to conduct the analysis.

9.2.2.3 Physical Abnormalities

Fish usability can be affected by altered appearance of fish. Obvious abnormalities may include:

- tumors and/or lesions on the body surface (including the eyes, lips, snout, gills, etc.),
- spinal column malformations,
- eroded, frayed, or hemorrhagic fins,
- other physical malformations,
- obvious parasites.

Any physical abnormalities should be reported for each fish, if present. For each class of abnormality that has been noted, a comparison between reference and exposure area fish should then be done using a chi-square goodness-of-fit test for relative frequencies. This information is used to help interpret effects, although for EEM purposes, a significant difference does not necessarily signify an effect. The number of statistical tests that are necessary will depend on the number of classes of abnormalities that are noted in the collected fish. Sample size will have been determined by the number of fish collected for the fish survey. Cohen (1988) provides guidance on the power of a chi-square test that would result from that level of replication.

9.2.3 Effects on Benthic Invertebrate Community

Table 9-5 and Table 9-6 outlines the four effect endpoints and appropriate statistics that are applicable for the benthic invertebrate community survey. In contrast to the fish survey, the statistical procedure used to determine whether there has been an effect is dependent on which of the six study designs is employed, as given in Table 9-6. For a given study, all four endpoints are analyzed using the same study-design-determined statistical procedure. The one exception is the Reference Condition Approach, which uses a different set of statistical procedures that do not require inter-area comparisons of these four endpoints, unless accompanied by ANOVA analyses; the procedures for this study design are outlined below and in Chapter 5.

Table 9-5: Effects on Benthic Invertebrate Community.

Effect endpoints to be used for determining statistically significant differences between exposure and reference areas or along an exposure gradient.

Effect Endpoint
Total invertebrate density
Taxon (e.g., family) richness
Simpson's Diversity Index
Bray-Curtis Index

Table 9-6: Statistical procedure used to determine an effect for each of the six study designs.

Study Design	Statistical Procedure (dependent on study design)
Control/Impact (C/I)	ANOVA
Multiple Control/Impact (MC/I)	ANOVA
Simple Gradient (SG)	Regression/ANOVA
Radial Gradient (RG)	Regression/ANOVA
Multiple Gradient (MG)	ANCOVA
Reference Condition Approach (RCA)	Multivariate/ANOVA

Note: Multivariate analyses can be performed on data collected using any of the designs in Table 9-6 to look for patterns (i.e., not hypothesis tests) that may be useful for highlighting potential areas of concern.

These designs fall into three basic categories with different philosophical approaches. These are:

1) the Control/Impact or Multiple Control/Impact designs, which are ANOVA type designs used to detect differences between discrete reference and exposure areas;

2) the Gradient (simple, radial or multiple) designs, which are intended to examine changes in community structure along a physical and/or effluent gradient which are better suited to regression analyses or ANCOVA;
3) the multivariate procedure of the Reference Condition Approach, which compares potentially "impaired" stations in exposure areas to a distribution of reference stations.

It should be noted that there may be some circumstances where ANOVA analyses are applicable to 2) and 3) above.

As for the fish survey, the endpoints in Table 9-5 should be reported in both graphical and tabular format for each area (reference and exposure area(s)) being summarized. The reported data should include means, medians, minimum and maximum values, standard deviations, standard errors, and sample sizes. Gradient data should be presented graphically as scatterplots of endpoint vs. distance from the effluent outfall. For gradient designs with no discrete "areas", tabular presentation prior to the main analysis would be applicable to station-by-station summary statistics, with the sampling unit being field subsamples rather than stations. Station-by-station summary statistics are also applicable to Control/Impact-type designs in cases where field subsamples are not pooled prior to taxa enumeration, although the key summary statistics are those used to determine effect and calculated for areas.

The same three main analysis steps outlined in Section 3.1 should be followed to determine whether statistically significant "effects" have occurred:

1) The data should be inspected to see whether they satisfy the assumptions of the statistical test or procedure being used (ANOVA, ANCOVA, Regression, Multivariate).
2) The appropriate statistical procedure (see below) would be performed following data inspection and any necessary transformations (or nonparametric alternative).
3) The key results for the effect endpoints (Table 9-5) should then be presented to clearly indicate whether there has been a statistical effect, with details on the nature of the effect (including direction and magnitude). Again, a statistical effect is declared if the *p* value exceeds the *a priori* α value determined as outlined in Section 4.1.

The same considerations and constraints discussed in Section 3.1 for conducting ANOVA and ANCOVA analyses apply to benthic invertebrate community analyses using those two statistical procedures. Thus, data inspection, analysis, and interpretation when using ANOVA or ANCOVA for the benthic invertebrate community survey should follow the generic recommendations given in Section 3.1. The following discussion reviews and highlights some of the additional considerations that arise when analyzing the different study designs that have been recommended for the benthic invertebrate community survey.

Although it is possible to use ANOVA to analyze data collected under most of the study designs listed in Table 9-6, ANOVA is most applicable to the Control/Impact (C/I) and Multiple Control/Impact (MC/I) designs. The simplest of these study designs is the C/I (or reference/exposure) design. In rivers, for example, this consists of one (usually upstream) reference area and one or more downstream exposure areas. Chapter 5 provides guidance on the different ways that C/I designs can be laid out. This kind of study design employs ANOVA comparisons between reference and exposure areas, with a significant difference signifying an effect.

The MC/I design is similar to the C/I design, except that it employs additional reference areas that are located in adjacent watersheds or bays where the sampled habitat is comparable to that found within the exposure area. This kind of design helps to avoid problems with confounding factors (e.g., when a single reference area differs from an exposure area with respect to several environmental variables in addition to the point-source effluent). Analogous to a C/I design, a significant difference between an exposure area and the mean of the reference areas, as determined by ANOVA, would represent an effect.

In contrast to the C/I and MC/I designs, the Simple and Radial Gradient designs are more amenable to regression analysis. The assumptions for regression analysis are applicable to the analysis of the benthic invertebrate community data, and have already been outlined in the Section 3.1 discussion on ANCOVA (regression is one component of ANCOVA). These assumptions are that:

- the residuals are normally and independently distributed with a mean of zero;
- the independent variable is fixed and measured without error; and,
- the relationship has the form specified (linear regression).

Gradient designs are particularly useful for 1) situations where rapid effluent dilution precludes the selection of an exposure area that is comparatively homogeneous in terms of effluent concentration and 2) determining how far along an effluent path effects are observed (i.e., determining the geographical extent of "effects"). The geographic extent of "effects" can be determined graphically by plotting the response variable(s) against distance from the effluent outfall and inspecting the data for an inflection point where the response variable asymptotes to the reference condition. Data from sampling stations arrayed in this manner could also be used, together with measured physicochemical data, in a multivariate analysis (e.g., ordination or clustering) used to identify which more distant stations tend to group with reference stations and which tend to group with clearly impacted stations.

Both of these approaches (graphical plotting and multivariate analysis) look for patterns in the data to qualitatively determine the approximate geographic extent of an effect. That is, they do not entail hypothesis testing and therefore, in the context of the EEM program, are not used to designate an effect sufficient to warrant follow-up action, but rather are used for informational purposes.

Nevertheless, statistical tests are possible for some gradients. In the simplest case, a statistically significant effect would be declared if the slope of the regression of an endpoint against distance from the effluent source is significantly different than zero (or if the correlation coefficient is statistically significant) (data transformations may be necessary to satisfy assumptions of linearity). In this case, the effect is a relatively uniform gradient of endpoint values away from the point source, rather than an effect in a given discrete area.

An effect can also be signified by a significant exposure vs. reference ANOVA difference when comparing a group of stations along the gradient close to the mine to "reference" stations along the gradient far from the mine. This is analogous to the C/I approach and assumes some degree of uniformity in exposure within the exposure group of stations and within the "reference" group of stations. Furthermore, the two groups of stations would need to be far enough apart to represent clear differences in exposure, and a sufficient number of stations would need to be available for each group to attain the desired level of power. Depending on the layout of stations, these assumptions may or may not be valid.

Given sufficient subsamples per station, it is also possible to use ANOVA to determine the presence or absence of an effect for a given station. This would entail using field-subsamples as replicates (treating stations as areas) and making station-by-station ANOVA comparisons of more near-field stations along the gradient to more distant reference stations. This method of analysis could be used to determine where along the gradient an effect disappears at the given α level of significance. This latter approach may, however, require extensive sampling effort, depending upon the number of stations along the gradient and the required (by power analysis) number of field sub-samples per station.

In cases where these kinds of statistical tests are not justifiable for a given gradient design, then a redesign of the monitoring program will be necessary to enable an appropriate statistical test during the next monitoring study. The re-design may entail increased replication focussed on the key exposure and reference areas (or stations) that are to be compared (e.g., increased replication in the near-field area of greatest effluent exposure and in the more far-field area that best represents reference conditions).

In some cases, it may be necessary to compare exposure versus reference gradients. This would be the case when a co-occurring (non-mine related) environmental gradient (i.e., covariate) confounds effluent effects in the exposure area. By using a multiple gradient design, it may be possible to make statistical comparisons of the exposure area gradient to a similar (non-mine related) environmental gradient in a nonexposed reference area. The reference gradient should be as similar as possible in depth and habitat to the exposure area gradient. Potential effluent "effects" would be tested for by using ANCOVA to compare reference to exposure area regression elevations (or adjusted means) while factoring out the influence of the co-occurring environmental covariate.

So, for example, if the gradient in effluent exposure away from the mine was confounded by a co-occurring increase in depth, an ANCOVA comparison might be made to a reference area where the depth gradient is the same. If the slopes for the reference and exposure area regressions against the covariate (X=depth) are approximately equal, then a significant difference in adjusted means would indicate an effect of the effluent on the endpoint Y (e.g., taxon richness). Again, Section 3.1 provides further guidance for ANCOVA analyses and the different ways these analyses can be used to indicate an effect.

The Reference Condition Approach (RCA) is a study design that combines inspection of multivariate patterns in the data with assessments of whether exposure stations fall outside a given confidence interval ellipse for reference stations. The fundamental concept of the RCA is to establish a database of stations that represent unimpaired conditions (reference stations) at which biological and environmental attributes are measured. This database is used to develop predictive models that match a set of environmental variables to biological conditions. These predictive models then allow a set of environmental measurements to be made at a new station and used in the model to predict the expected biological condition at the new station. A comparison of the actual biological condition at the new (exposure area) station with conditions at the reference stations to which the new station is predicted as belonging allows an assessment of whether there has been an effect at the exposure station.

The reference condition database is established by an initial standardized sampling program at a wide variety of spatial scales. The same benthic macroinvertebrate sampling protocol is used in as many ecoregions and stream orders or lakes as are available in a catchment. A number of environmental variables are measured in conjunction with invertebrate sampling. The data are then subjected to a 3-step multivariate analysis in which:

1) a number of invertebrate groups are formed based on similarity of community structure;
2) biological data are correlated with environmental attributes, and an optimal set of environmental variables is identified that can be used to predict group membership;
3) the biological condition of test (exposure) stations is assessed by using the optimal set of environmental variables to predict group membership.

How the test station fits, relative to the group to which it is predicted to belong, establishes whether, and to what degree, the station is different from the reference group. If a station or group of stations fall outside the statistically-determined, confidence interval ellipse for the reference stations, this signifies the presence of an effect. The boundaries of the reference ellipse should be set *a priori* based on some of the considerations discussed in Section 4.1. A more complete discussion of the assumptions, procedures, and interpretation of the RCA is available in (Reynoldson et al., 1995, 2000). It should further be noted that, depending on the timing and locations of an RCA sampling program, it may also be possible to use the resulting database to make ANOVA comparisons between reference and exposure areas to determine whether there has been an effect. This latter kind of analysis would be analogous to a MC/I design.

To summarize, an overall procedure similar to that outlined at the end of Section 3.1 should also be followed (with appropriate modifications) for the benthic invertebrate community survey. To reiterate, the procedure should include:

1) preparing the analyses: QA/QC (including checks for data entry errors), summary of confounding factors, description of the sampling design and taxonomic level used, clear identification of the sampling units used for statistical comparisons (e.g., stations rather than field-subsamples), ensuring equivalence of sampling substrata and sampling techniques among different reference and exposure areas being compared;
2) initial summary statistics (graphical and tabular presentation of means, etc., as described above);
3) hypothesis-testing statistical analyses to determine "effects" (ANOVA, ANCOVA, regression);
4) graphical approaches (e.g., inspection of the shape of regression lines; used for inspecting patterns in the data rather than determining "effects");
5) multivariate statistical analyses used for determining a) patterns in the data and b) the position in multivariate space of exposure stations relative to reference confidence ellipses; only b) used to determine "effects";
6) power analyses (not applicable to graphical approaches and RCA).

The following benthic invertebrate community endpoints should also be reported, including means, standard deviations, minimum and maximum values, and sample sizes:

- Evenness,
- Taxon (e.g., family) density,
- Taxon (e.g., family) proportion,
- Taxon (e.g., family) presence/absence.

Unlike those in Table 9-5, these endpoints are included as supporting variables, and are not statistically analyzed to determine "effects". They may, however, be used to interpret effects at later stages (e.g., determining the magnitude and causes of "effects"). These should be reported in both graphical and tabular format for each area (reference and exposure area(s)) being summarized. It should be noted that there may be other descriptors that may also be useful for the interpretation of monitoring data, on a site-specific basis (see Resh et al., 1995 for a review).

9.3 The Role of Power Analysis, α, β, and Effect Size in Determining Effects

9.3.1 Setting αand β

There are always potential errors in any statistical evaluations of "effects". In testing whether exposure areas differ significantly from reference areas, a low probability of a Type I error (α) is usually allowed so that a normal population or community will not be mistaken for an impacted one. However, the monitoring program should also be designed to provide a reasonably high probability of statistically detecting a predetermined effect size (ES) if it has occurred. In other words, the power of the test should be high. Power is 1-β, where β is the Type II error (see below).

Type I error is partially kept in check by setting a broad margin for variation around what is considered "healthy". This is reflected in the methods used for setting ES discussed in the Guidance Document. Sufficient sampling effort should also be expended to reduce Type II error, taking into account the low probability allowed for Type I error. Thus, to determine what sampling effort is required, ES, Type I, and Type II error will all be taken into account and set *a priori*.

Decisions should be made about the magnitude of Type I and Type II errors that are acceptable for determining power and thus the sampling effort required to detect the recommended ES.

Type I error occurs (at probability α) if the null hypothesis that there is no effect is rejected when in fact it is true (e.g., an exposure area is declared as being different from reference when it is not).

Type II error occurs (at probability β) if the null hypothesis is accepted when it is false (e.g., the exposure area is declared as not being significantly different from reference when it is actually impaired). Therefore, α is the risk to industry, and β is the risk to the environment.

The power of a statistical test is 1-β, the probability associated with correctly rejecting the null hypothesis when it is false (e.g., the probability associated with correctly identifying an impaired area). In a well designed, properly replicated monitoring program, the goal is to keep α and β low and power high.

As can be seen from the equation given later in this section, one way to increase power, given a fixed sampling effort (i.e., sample size), is to increase α. In other words, there are trade-off decisions to be made when setting α and β. Traditionally, α has been set at

0.05 for experimental studies where, in many cases, the cost of a Type II error is not particularly high. That is, an α of 0.05 is typically used in situations where the primary concern is to have maximal confidence that a statistically significant effect is real. On the other hand, there is much less consensus and available literature on what is an appropriate level for β. Some studies have suggested using a minimal power of 0.8 (i.e., β = 0.2; Alldredge, 1987; Cohen, 1988; Burd et al.,1990; Osenberg et al.,1994; Keough and Mapstone, 1995).

In many cases, this "rule of thumb" can be traced back to Cohen's seminal work on power analysis (see Cohen, 1988), which is primarily geared toward applications in the behavioral sciences. For those kinds of applications, Cohen argued that Type I errors were likely to be more serious than Type II errors, for cases where the biggest concern is to not propagate erroneous conclusions based on incorrect declarations of significant differences. Specifically, he suggested that, if Type I errors were to be considered four times more serious, then it might be reasonable to set α at the traditional (in terms of experimental studies) 0.05 and β at: 4 x 0.05 = 0.2. He specifically warned, however, that this rule of thumb should be ignored for other kinds of studies where these assumptions are not applicable.

This latter caveat applies to environmental monitoring studies where, because of the potentially high cost (both ecological and monetary) of failing to detect negative impacts, many researchers in the field of biomonitoring argue that α should be set at least to the same level as β (e.g., Alldredge 1987; Underwood 1993; Mapstone 1995). In other words, the argument has been widely made that, barring extenuating circumstances, the risk to the environment should not be set greater than that to industry. This suggests that the most reasonable starting point is to set α = β, and this position has been adopted by the EEM program. On a site-specific basis, it may sometimes be decided to 1) set α > β if it can be shown that the risk to the Environment is of greater concern than the risk to Industry or to 2) set α < β if it can be shown that the risk to Industry is of greater concern.

After deciding to set α = β, it is then necessary to make a decision on an appropriate value for α and β. In many cases, this decision will be made within the context of the desired power of the test, the ES that the program is to be designed to detect, and the implications for sampling effort. This decision making process can be illustrated with an example from Chapter 5. In this example, the effects on sample size of setting α and β at different levels were examined for detecting an effect of ±2 SD (Table 9-7) by using the following power analysis equation, which yields an approximate sample size (n) in one step for the most basic Control/Impact ANOVA design (see also the discussion in the next section for further details) (Guenther 1981; Alldredge 1987; Environment Canada and Department of Fisheries and Oceans 1995):

$$n = (2(Z_a + Z_b)^2(SD/ES)^2) + 0.25(Z_a)^2$$

where:

n = sample size
Z_a = standard normal deviate for α significance level (Type I error)
Z_b = standard normal deviate for β significance level (Type II error)
SD = standard deviation
ES = effect size

Table 9-7: Sample sizes required to detect difference of ±2 SD

α	1-β			
	0.99	**0.95**	**0.90**	**0.80**
0.01	14	11	10	8
0.05	11	8	7	5
0.10	9	7	5	4

SD, α, and β as above

Using Table 9-7 for guidance (and the recommendation that the benthic invertebrate community survey should minimally have sufficient power to detect an ES of ±2 SD), the benthic invertebrate working group recommended α and β be initially set at 0.1. This implied that, in most cases, the sampling effort would require a sample size of 5, which is within the range used in many benthic surveys (Resh and McElvray, 1993). Basic ANOVA power analysis calculations also indicate that α and β can be set equal to 0.1 for the fish survey endpoints, as well, with very little effect (relative to $\alpha=0.05$, $\beta=0.2$) on the sample size required to achieve the resulting level of power (1-β). The use of an α or β level other than 0.1 would require appropriate justification by either the proponent or the AO (e.g., setting α more rigorous, lower Type II error (β) when the risk to the environment is judged to be of greater concern). Consultation with the AO may also be required in cases where power analysis recommends the use of unreasonably high sample sizes.

It should also be noted from Table 9-7 that, by increasing sample size, it is possible to obtain lower Type I and II errors (lower α and β), while still maintaining α equal to β. For example, α and β can both be set at 0.05, resulting in 95% power to detect an ES of ±2 SD, by increasing sample size to 8 (see Table 9-7). The same argument applies to the other components of EEM (e.g., the fish survey and fish usability components) for different desired ES's, although the required sample sizes will be different. Thus, setting α equal to β provides an economic incentive to carrying out a well designed, well replicated monitoring program, because providing sufficient replication will help to reduce the probability of Type I errors (i.e., α is kept low), thereby reducing the probability of unnecessary follow-up studies. Furthermore, since α is linked to β, the power of the monitoring program to detect real effects will also be increased. This improvement in monitoring design helps to ensure a better understanding of what kinds of effects, if any, are occurring.

9.3.2 Power Analysis: Determination of Required Sample Size, Power, and Appropriate Effect Size

Power analysis is used for two major purposes during EEM:

1) at the beginning of a monitoring study, to calculate the sampling effort (sample size) that will be required to detect a given ES at a given level of power;
2) following a recently completed monitoring study, to determine the level of power that was actually achieved.

Both of these uses of power analysis are briefly reviewed here to help clarify the relationship between the two.

During the initial design phase of an EEM study, power analysis can be used to determine the sample size required to achieve a test adequate to detect an effect with a magnitude = ES, where ES has been determined prior to sampling. Using ES, α, β and some estimate of reference variability (e.g., standard deviation, SD), and making some assumptions about the distribution of the data being evaluated, a scientifically defensible sampling strategy can be devised. The discussion below outlines the most basic (i.e., Control/Impact ANOVA) procedure for determining required sample size. Sample size refers to number of fish for the fish survey and number of stations for the benthic invertebrate community survey. In cases where the required sample size calculated for one endpoint (e.g., invertebrate density) is greater than that calculated for another (e.g., invertebrate taxon richness), the greater sample size should be used (unless, as discussed above, consultation with the AO confirms that this would result in excessively high sample sizes).

Once ES has been determined, the levels of α and β have been selected, and SD for the particular mine location in question has been estimated, they would then be plugged into the power analysis equation to calculate the sample size required to detect an impact of magnitude ES between or among areas at a given power level. For the most basic ES Option in Chapter 5 (ES = ±2SD), due to canceling of terms, determination of SD is not required for the power analysis and Table 9-7 above gives precalculated sample sizes for various values of α and β.

It should be noted that determination of required sample size assumes that the variability amongst replicates for the exposure area is similar to that for the reference area. Although ANOVA's are fairly robust with respect to violation of normality assumptions, if the variance within an exposure area is much higher (or lower) than within the reference area, ANOVA comparisons may not be appropriate, unless the variances can be made homogeneous by transformation. For the case where the exposure and reference variances remain significantly different following transformation, the power analysis outlined here may over- or more likely under-estimate the number of sampling stations required. Non-parametric tests may be used; however, the estimate of power described herein will no longer be relevant. In some cases, unequal variances may be an artifact of insufficient field sub-sampling. See the benthic invertebrate community Guidance Document for further guidance.

For a basic Control/Impact ANOVA design, the estimated sample size (number of replicate stations per reference and exposure area) required to detect a given ES at a given power level can be calculated by arranging the standard power analysis equation as follows (Green 1989):

$$n = 2(t_a + t_b)^2(SD/ES)^2$$

where:

n = sample size
t_a = t value for αsignificance level (Type I error; two-tailed)
t_b = t value for β significance level (Type II error; one-tailed)
SD = standard deviation
ES = effect size

The equation is solved iteratively by choosing an approximate value of n to look up t_a and t_b and then using the solution to find a more accurate n; the procedure is repeated until arriving at a final estimate for n. Alternatively, the equation given in Section 4.1 can be used to approximately solve for n in one step. Pre-calculated tables of n (expanding upon Table 9-7) are available for a variety of values of α, β, and ES (Alldredge, 1987; Cohen, 1988; Environment Canada and Department of Fisheries and Oceans, 1995).

The reader is referred to the appropriate literature (e.g., Cohen 1988) for guidance on power analysis and tables for determining sample size for regression, ANCOVA, and chi-square monitoring designs. As for a basic Control/Impact design, power analysis for these other designs will also require an *a priori* decision on an appropriate magnitude for ES. As discussed below, this is likely to be an evolving process. For regression analyses, Cohen (1988) gives a table for converting ES's from SD units to a correlation coefficient (r), and in some cases, it may be reasonable to use this r to look up the approximate sample size required for a regression-type gradient design. For example, given certain assumptions, he shows that using an ES of 2 SD is equivalent to using $r = 0.707$ (or $r^2 = 0.5$). Although the exact equivalency depends on the assumptions involved, it may be acceptable to use this conversion (possibly with a correction factor) to obtain an approximate ES appropriate for use in regression-type analyses. Tables are provided in Cohen (1988) for looking up required sample sizes for various values of r, α and β.

In addition to using power analysis at the beginning of a study to calculate the sample size required to detect a predetermined ES, the above equation can also be rearranged to calculate the power that was achieved in a study that has just been completed. This is done by solving for β rather than n (power = $1-\beta$). The value used for α is that used to make the statistical comparisons between reference and exposure areas (as outlined in Sections 3 and 4.1). ES should be based on the most current estimate of the magnitude of an effect that the EEM program would like to be able to detect (see below). The power analysis will also include the sample size used during the completed study and the calculated standard deviation (or equivalent variability measure, such as the square root of the MSE from an ANOVA or ANCOVA) used to perform the statistical comparisons among reference and exposure areas. This power analysis at the end of a study is particularly important if there has been a change in any of the relevant parameters that could affect power ($1-\beta$) and that were used or calculated at the beginning of the study (i.e., α, ES, n, SD). If none of these parameters (or anticipated statistical procedures) have changed, then the post-study power analysis would likely be a review of the power analysis performed at the beginning of the study. It should also be noted that it may sometimes be useful to rearrange the power analysis equation to calculate the ES that the study would be expected to be able to detect, given the sample size employed and the designated α and β levels (i.e., solve the equation for ES). Finally, the information gained by performing post-study power analyses should also be used to provide guidance for study design (particularly sampling effort) for the next monitoring study.

The fish (Chapter 4) and benthic invertebrate community (Chapter 5) chapters include guidance on how to set appropriate ES's for each kind of study design. Given the current shortage of information on appropriate ES's (see Lowell, 1997), determination of these ES's is likely to be an evolving process. Consequently, the ES's currently in use should be reevaluated at the beginning and end of each monitoring study and ES's should be readjusted as necessary. Furthermore, when evaluating temporal trends in "effects", careful consideration should be given to the ES's used to determine sample sizes during past monitoring studies. In some cases, it may be determined that replication during past monitoring studies was insufficient to provide adequate power to detect the newly readjusted ES's. In those cases, the (newly determined) low power employed during past monitoring studies may have resulted in missing important previous, and perhaps more subtle, "effects".

To avoid confusion, it should be noted that the term "effect size" is sometimes variously used to refer to three different concepts:

1) The first is the magnitude of an effect that is actually measured in a given study (e.g., a given study may measure a 10% decrease in relative gonad size downstream of an effluent outfall, relative to a reference area).
2) The second concept can be termed a "critical effect size"; it is a predetermined value and refers to what would be considered to be an effect of sufficient magnitude to be of ecological, economic, or social significance.
3) The third kind of "effect size" is that used in this document (ES in the above equation) to calculate required sampling effort and to determine statistical power; it is also predetermined.

As already noted, the ES used here to determine sampling effort does not necessarily imply that an effect of that magnitude is of ecological, economic, or social significance. Rather, it is set based on the magnitude of an effect that the EEM program would like to be able to detect. This magnitude will likely be at least partly based on *initial* ideas of what would be an ecologically (or other) important effect. But it may later be decided that greater or lesser effects are of interest with respect to taking corrective action. For example, ES for fish relative gonad size may be set at ±25% at the beginning of a study for the purposes of determining sampling effort. But following the study, it may be decided that a 10% effect is unacceptable due to priorities set by local stakeholders. In the same vein, it may later be decided that a 40% effect within a prescribed area may be acceptable due to local economic concerns. Thus, ES used during a study is related to, but does not necessarily equal, a "critical effect size" used for making decisions regarding the need for corrective action.

9.4 Alternative Methodologies for the Fish and Benthic Invertebrate Community Surveys

For Metal Mining EEM, 1) caged bivalve on-site bioassays (fish survey), 2) expanded sublethal toxicity testing (fish survey), and 3) mesocosm on-site bioassays (both fish and benthic invertebrate community surveys) have been recommended as potential alternative methodologies for monitoring when confounding influences or difficult field conditions preclude the usefulness of the standard field surveys. See Chapter 12 for more details on alternative methods for monitoring.

The mesocosm documentation provides details for application to EEM studies, but some generic data interpretation guidance can be provided here. The most likely design for an EEM mesocosm study would be the equivalent of a reference area versus near-field exposure area ANOVA comparison. That is, endpoints measured in replicated mesocosms mimicking reference conditions would be compared to the same endpoints measured in replicated mesocosms mimicking near-field conditions.

For the fish survey, it would be necessary to choose endpoints with a rapid enough response time that any potential "effects" would be picked up during the length of time that the fish are exposed in the mesocosms. Some examples of effect endpoints that could be used include individual growth, condition factor, liver size, gonad size, and time to maturity. In order to detect changes in these endpoints over the length of a typical mesocosm exposure period (e.g., 30 days), it may be necessary to follow individual fish (e.g., for growth) or groups of fish that all began at approximately the same size (e.g., for liver size). The particular endpoints that are chosen will depend on the requirements and constraints of each individual study.

If the fish are chosen so that they are all within a comparatively narrow size/age range, then it would not be necessary to use ANCOVA to factor out the influence of the size/age covariate on the endpoints. In this case, simple ANOVA comparisons could be used for all the endpoints. If limited fish availability necessitated the use of a wider size/age range, then ANCOVA comparisons might be required for some of the endpoints to factor out size age/effects, as described in Section 3.1. In either case, the same data analysis procedures outlined in Section 3.1 would apply.

Thus, the three steps of the main statistical analyses would be to:

1) Inspect the data to see that the assumptions of ANOVA or ANCOVA are met and carry out any necessary data transformations.
2) Conduct the appropriate statistical comparisons (including nonparametric alternatives, if necessary).
3) Clearly present whether or not there have been any statistically significant "effects" and, if so, indicate the direction and magnitude of the "effects". As for the field survey, significance would be judged against an a value chosen *a priori* as discussed in Section 4.1.

And again, as outlined in Section 3.1, the same overall steps of data preparation, presentation, and analysis would apply:

1) preparing the analyses,
2) initial summary statistics,
3) ANOVA and/or ANCOVA analyses (include the three main statistical analysis steps outlined above),
4) power analyses.

It should be noted that it may be possible to use the fish from a fish survey mesocosm study for the fish usability component of EEM, as well, provided that the effect endpoint has the potential for responding over exposure periods typical of mesocosm studies. In this case, the endpoints and statistical procedures would be as outlined in Section 3.2.

Some considerations would be unique to a mesocosm-type study. For example, control over experimental considerations would likely result in lower levels of variability within reference and exposure treatments, as compared with field data. This may make it possible to attain equivalent levels of statistical power using smaller sample sizes than used in the field. In the same vein, it may be possible to attain higher power levels or to detect smaller effect sizes while using the same sample sizes as used in the field. And, in fact, it may be desirable to have sufficient power to detect smaller effect sizes in mesocosm studies than in field surveys, due to the shorter exposure times typical of mesocosm studies. That is (using hypothetical numbers), a 10% effluent-induced change over a 30 day exposure period in a mesocosm study may be equivalent to a 25% change over a much longer lifetime exposure in the field.

In addition, due to the possibility of caging artifacts, it may be necessary to switch from using individual fish as the sampling unit for replication (as in the field) to using individual experimental enclosures (mesocosms) as sampling units. In other words, simply using two mesocosm units (one for reference and one for exposure) with 20 fish each may not be valid because it may not be possible to separate effluent effects from the effects due to subtle differences in the experimental enclosures themselves. This is an example of the potential for confounding effects due to "pseudoreplication" (Hurlbert, 1984).

In comparison to the fish survey, it may be even more straightforward to substitute mesocosm studies for benthic invertebrate community field monitoring, at least in terms of statistical design and analysis. As for the fish survey, the same steps outlined in Section 3.3 would apply for data preparation, presentation, and analysis. Furthermore, due to comparatively fast turn-around times for changes in invertebrate community structure within mesocosms, it may be possible to use the same effect endpoints as used in the invertebrate field survey (Table 9-5). The most likely study design would be analogous to the Control/Impact design (Table 9-6), with ANOVA comparisons being made between replicated reference and exposure mesocosms. The sampling units would be the individual mesocosms (equivalent to "stations" in the field survey). As for fish mesocosms, control over variability under experimental conditions may make it possible to attain greater statistical power or to detect smaller effect sizes (in terms of % change) using the same sample sizes as typically used in the field. This increase in *precision* is one of the most frequently cited advantages of using mesocosms in place of field sampling, and is weighed against the disadvantage of a potential decrease in *accuracy* due to using a (hopefully realistic) simulation of actual field conditions.

9.5 Sublethal Toxicity

Two major uses of the results of sublethal toxicity tests are to track deviations in effluent quality and to understand the relative contribution of a mine's effluent in multiple discharge situations. These uses are not discussed further here, other than to say that sublethal toxicity endpoints (IC_{25}, EC_{25} and LC_{50}) should be tabulated as geometric means and confidence intervals and graphically presented to show trends. Further, sublethal toxicity results should be qualitatively compared to the respective field components to see if any relationships between field and lab tests exist.

Given their emphasis on effluent characterization, this Section does not list effect endpoints for sublethal toxicity tests as done in Tables 1-3 for the fish and invertebrate surveys and for fish usability. Nevertheless, in the absence of measured "effects" for the given fish and invertebrate endpoints, data from the sublethal toxicity tests may signal the need for further monitoring, although this is not a requirement. For example, sublethal algal tests may indicate the potential for toxicity to algal communities, and these effects may not be picked up in the fish or invertebrate monitoring programs. In this case, the results from sublethal toxicity testing may indicate the need for additional monitoring. Similar arguments regarding potential indicators of the need for further monitoring could be made for the use of water and sediment quality data.

9.6 References

Alldredge, J.R. 1987. Sample size for monitoring of toxic chemical sites. Environ. Monit. Assess. 9:143-154.

Burd, B.J., A. Nemec, R.O. Brinkhurst. 1990. The development and application of analytical methods in benthic marine infaunal studies. Adv. Mar. Biol. 26: 169-247.

Cohen, J. 1988. Statistical power analysis for behavioral sciences, 2nd Ed. Lawrence Erlbaum Associates, Hillsdale, NJ.

Draper, N.R. and H. Smith. 1981. Applied regression analysis. 2nd ed. John Wiley & Sons, New York.

Environment Canada. 2001. Guidance document for determining follow up actions when effects have been identified in environmental effects monitoring (EEM). Draft document. Hull, QU.

Environment Canada. 1998. Pulp and paper technical guidance document for aquatic environmental effects monitoring. EEM/1998/1.

Environment Canada and Department of Fisheries and Oceans (EC/DFO) 1995. Further guidance for the adult fish survey for aquatic environmental effects monitoring related to federal *Fisheries Act* requirements. EEM 1. Environment Canada, Ottawa, Ont. 72 pp.

Guenther, W.C. 1981. Sample size formulas for normal theory *T* tests. Am. Stat. 35: 243-244.

International Standards Organization (ISO). 1992. Sensory Analysis: Vocabulary. International Standards Organization, Geneva, Switzerland. ISO 5492.

Keough, M.J. and B.D. Mapstone. 1995. Protocols for designing marine ecological monitoring associated with BEK mills. Technical Report Series No. 11, National Pulp Mills Research Program, CSIRO, Canberra, AU.

Lowell, R.B. 1997. Discussion paper on critical effect size guidelines for EEM using benthic invertebrate communities. EEM/1997/8. Environment Canada, Hull, QU.

Mapstone, B.D. 1995. Scalable decision rules for environmental impact studies: effect size, Type I and Type II errors. Ecol. Appl. 5:401-410.

Osenberg, C.W., R.J. Schmitt, S.J. Holbrook, K.E. Abu-Saba and A.R. Flegal. 1994. Detection of environmental impacts: natural variability, effect size and power analysis. Ecol. Appl. 4: 16-20.

Peters, R.H. 1983. The ecological implication of body size. Cambridge University Press, New York. 329 pp.

Resh, V.H. and E.P. McElravy. 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates, p. 159-194. In Rosenberg, D.M. and V.H. Resh. (ed.) Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

Reynoldson, T.B., K.E. Day and T. Pascoe. 2000. The development of the BEAST: a predictive approach for assessing sediment quality in the North American Great Lakes, p. 165-180. *In* Wright, J.F., D.W. Sutcliffe and M.T. Furse (ed.), Assessing the biological quality of fresh waters. RIVPACS and other techniques. Freshwater Biological Assoc. Ambleside, U.K.

Reynoldson, T.B., R.C. Bailey, K.E. Day, R.H. Norris. 1995. Biological guidelines for freshwater sediment based on benthic assessment of sediment (the BEAST) using a multivariate approach for predicting biological state. Austral. J. Ecol. 20: 198-219.

Sokal, R.R. and F.J. Rohlf. 1995. Biometry, 3rd Ed. W.H. Freeman, New York.

Underwood, A.J. 1993. The mechanics of spatially replicated sampling programmes to detect environmental impacts in a variable world. Austral. J. Ecol. 18:99-116.

Appendix 1.1 Summary Questions for Fish Survey

The following items provide guidance for some of the key questions that should be addressed when summarizing the results of a particular fish survey, using Pulp and Paper Cycle 2 as an example.

1. List possible confounding factors and the geographic extent of their effects.
2. Was an alternative to the standard fish survey used?
3. What species were used?
4. If these are different from cycle 1, why were they changed?
5. Were sample sizes based on power analysis? Give the original standard deviations, power levels, etc.
6. How many were collected; what methods were used; how much gear; how many days of effort? What was the CPUE (if recorded)?
7. Were data collected from > 1 exposure or reference area?
8. Are there problems with the reference area?
9. Is there doubt about movement between areas?
10. Was a tracer used?
11. Are there statistically significant differences between areas that increase confidence that fish are not mixing?
12. Were immature fish included in analyses for adult fish (note: for most species a GSI < 1% represents an immature fish)?
13. Is there evidence that either of the species may have adults which spawn in alternate years (e.g., GSI > 1%, but adult GSI's are bi-modal)?
14. Can power be calculated for cycle 3?

ANOVA ANALYSES

1. Was there a statistically significant difference in age for either species or sex (*p* value; % difference; which mean was larger)?
2. Was there a statistically significant difference in weight?
3. Was there a statistically significant difference in length?
4. Would including immature fish in the analysis improve interpretability?
5. Can any information be extracted from the data on the fish collected re: size or age at maturity?

ANCOVA ANALYSES

1. For condition factor - were there significant regressions?
   - were the slopes approximately equal between areas (i.e., lack of significant interaction term for covariate vs. area)?

   If yes, did plots of the data provide interpretable data? If the slopes were equal, was there a significant difference between areas? If yes, which area had the larger adjusted mean? What was the % difference in slope(if slopes significantly different)? What were the *p* values for the slope and adjusted mean comparisons?

2. Same questions for size-at-age, gonad size, liver size, egg size, fecundity.

Appendix 1.2 Summary Spreadsheet for Fish Survey

This spreadsheet provides brief hypothetical examples of some of the key results that might be used to summarize the findings of a fish survey, using Pulp and Paper Cycle 2 as an example.

GENERAL INFORMATION	Mill 1
Was fish survey conducted?	yes
If no fish survey, why not?	NA
Was alternative to fish survey used?	no
Confounding factors (describe)	none
Cycle 1 sentinel species	tomcod
Cycle 2 sentinel species	mummichug
Reference area location for fish	(see maps)
Tracer used?	sum dehydroabietic, isopimaric & abietic acids in exposure areas were 1.3 x higher for males and 1.8x higher for females *not replicated or tested statistically*
Any movements between reference and exposure areas?	unlikely (see text)

FISH CAPTURE	
Target Sample Size (power analysis)	for Cycle III, power analysis recommends: - 20 males (for detecting a 10% difference in gonad weight) - 40-50 females (for detecting a 20-30 % difference in ovary weight)
Collection method	baited minnow traps & seine nets
Method different from Cycle 1?	yes, added seine nets
Date(s) of collection	May 18 & 19 (exposure) May 20 & 21 (reference)
Number of fish	exposure - 26 females, 35 males, 1 immature reference - 22 females, 25 males
CPUE	8.6
Do immatures affect statistical	no

calculations?	
Aging procedure	dorsal scales & otoliths
Fractional spawner?	yes
Problems with capture?	baited traps unable to collect mummichog; captures only at high tide

FISH RESULTS

Sex	female
Age: p value	0.006
Age: % difference	26% younger in exposure area
Weight-at-age: slopes significantly different than 0?	yes (reference - p=0.023) (exposure - p=0.011)
Weight-at-age: significant interaction?	no, p=0.716
Weight-at-age: p value for adjusted means	p= 0.812
Weight-at-age: % difference in adjusted means	not significantly different
Condition factor: slopes significantly different than 0?	yes (reference - p=0.001) (exposure - p=0.001)
Condition factor: significant interaction?	yes, p=0.0486
Condition factor: % difference in slopes	8.54% steeper slope in exposure area
Relative liver weight: slopes significantly different than 0?	no (reference - p=0.681) (exposure - p=0.552) (body weight as covariate)
Liver weight (without covariate): p value	0.524
Relative gonad weight: slopes significantly different than 0?	no (reference - p=0.097) (exposure - p=0.227 (body weight as covariate)
Gonad weight (without covariate): p value	0.229
What was the magnitude of change in: reference mean age minus exposure mean age, from Cycle 1 to Cycle 2 (i.e., temporal change)?	25% less age difference in Cycle 2
Did the plots provide interpretable data?	yes

Appendix 2.1 Summary Questions for Benthic Invertebrate Community Survey

The following items provide guidance for some of the key questions that should be addressed when summarizing the results of a particular benthic invertebrate community survey, using Pulp and Paper Cycle 2 as an example.

1. Identify possible confounding factors and the geographic extent of their effects.
2. Was an alternative to the standard invertebrate survey used?
3. What endpoints or summary variables were use to identify effects?
4. What taxonomic level was used?
5. At what stages was QA/QC used?
    - Taxonomy
    - Sub-sampling
    - Sorting
    - Data entry
6. What study design was used and did this change from Cycle 1?
7. Does the study design need further modification for Cycle 3?
8. How many reference and exposure areas were sampled?
9. Were multiple habitats sampled (e.g., depositional vs. erosional)?
10. What environmental variables were measured (e.g., freshwater - grain size, TOC; marine - Redox, C/N ratio)?
11. What was the scale of replication used (field sub-sample or station) and did this change from Cycle 1?
12. Were sample sizes based on previous power analyses?
13. What sampling device was used and did this change from Cycle 1?
14. What mesh size was used and did this change from Cycle 1?
15. What sub-sampling methods were used and did they change from Cycle 1?
16. What data analysis methods were used?
17. Was the interpretation based on quantitative or qualitative differences, or both (e.g., ANOVA vs. graphical regression)?
18. What endpoints showed an effect associated with exposure and what was the nature of the change (e.g., direction and magnitude of change of endpoint)?
19. Can the effect be partially or wholly ascribed to the effluent?
20. If there were no statistically significant effects, was the power sufficient?
21. Can power (and required sampling effort) be calculated for Cycle 3?
22. Was there a difference in the response associated with effluent exposure from Cycle 1 to Cycle 2?

Appendix 2.2 Summary Spreadsheet for Benthic Invertebrate Community Survey

This spreadsheet provides brief hypothetical examples of some of the key results that might be used to summarize the findings of a benthic invertebrate survey, using Pulp and Paper Cycle 2 as an example.

Endpoints used in designating effects	**Mill 1**	**Mill 2**
Abundance	yes	yes
Richness	yes	yes
Simpson's	yes	no
Bray-Curtis	yes	no
Other (what)	EPT taxa	Shannon-Weaver, evenness
Taxonomic level used		
Lowest possible	yes	no
Family	no	family for most
Higher than family	no	nematodes to phyla
QA/QC		
Taxonomic identification	yes	yes
Lab. sub-sampling	yes	yes
Sample sorting	yes (95% recovery)	yes (98% recovery)
Data entry	yes	yes
Study Design selected		
Control/Impact	yes	no
Multiple Control Impact	no	no
Simple gradient	no	yes
Radial gradient	no	no
Multiple gradient	no	no
Reference Condition Approach (RCA)	no	no
Other (what?)	no	no
Areas sampled		
Number of reference areas	2	NA
Number of exposure areas	2	NA
Replication within areas		
Number reference stations/habitat	5 (erosional) 5 (depositional)	NA
Number exposure stations/habitat	5 (near-field, erosional) 5 (far-field, depositional)	NA
Number of stations per gradient		15 (depositional)
Habitats sampled		
River – erosional	yes	no
River – depositional	yes	no
Lake – erosional	no	no
Lake – depositional	no	yes
Intertidal – hard	no	no
Intertidal – soft	no	no
Subtidal – hard	no	no
Subtidal –soft	no	no

Estuary	no	no
Proportional to occurrence	no	no
Other	no	no
Formation of receiving water body		
natural/man-made	natural	natural
Habitat variables		
Which measured?	grain size, TOC, velocity	grain size, TOC
Field sub-samples		
Date sampled	2-5 Oct., 1998	15-20 Sept., 1999
Number sub-samples per station	3	4
Pooled?	yes	no
Sampling device used (which)	Neill-Hess, Ponar	Ekman
Mesh size used (mm)	500	500
Lab. sub-sampling		
Sub-sampling method	Marchant box	cone sub-sampler
Data analysis methods		
ANOVA	yes	no
ANCOVA	no	no
Regression	no	yes (attempted)
Other	reciprocal averaging	principal components
Interpretation		
Presence of effect determined quantitatively?	yes (ANOVA)	not able to run analysis; see Comments below
What was effect ascribed to?	combined mill and municipal effluent	NA
Presence of effect determined qualitatively?	yes (reciprocal averaging)	yes (scatterplot)
What was effect ascribed to?	combined mill and municipal effluent	mill effluent
Which endpoints responded?	66% greater abundance in near-field area (ANOVA)	up to 70% lesser abundance with increasing exposure (scatterplot)
	altered taxa assemblages in near-field area (reciprocal averaging)	up to 46% lesser richness with increasing exposure (scatterplot)
Sample size selection from power analyses	yes	no
Was power sufficient?	yes	NA
Can power be calculated?	yes	yes

Was there a difference from cycle 1?	yes	yes
What was the difference from cycle 1?	less evidence of enrichment	more evidence of toxicity
Comments (e.g., confounding factors)	alternative methodology needed to separate mill and municipal effluent effects	appears to be nonlinear gradient of effects; field survey needs redesign to enable quantitative determination of effects

Table of Contents

10. Public Involvement in Metal Mining EEM *10-1*

10.1 Overview *10-1*

10.2 Objective and Potential Scope of Public Involvement in EEM *10-2*

10.3 Mechanisms for Public Involvement *10-3*
10.3.1 Ongoing Mechanisms 10-4
10.3.2 Other Mechanisms 10-5

10.4 Public Involvement Plan *10-6*
10.4.1 Provision of EEM Information to the Public 10-6
10.4.2 Proposed Mechanisms for Public Involvement 10-7
10.4.3 Proposed Mechanisms to Respond to Public Input 10-11

10. Public Involvement in Metal Mining EEM

10.1 Overview

The objective of this guidance is to help facilitate public involvement in the EEM program, particularly on a site-specific basis. As part of the EEM program mines are strongly encouraged to provide opportunities for public involvement in all aspects of the EEM program. Public input can play an important role throughout the EEM program, from the early planning steps prior to the initiation of EEM, to the preparation of the site characterization and the first study design, as well as the and data interpretation for each EEM study at a site, and decisions regarding next steps in the EEM program at a site.

Following recommendations of the Whitehorse Mining Initiative and AQUAMIN, public involvement is an important component of the Metal Mining EEM program. The Whitehorse Mining Initiative recommended the following as an underlying principle for public involvement: "more effective approaches to environmental management can be developed, and the public trust in mining enhanced, when the public and other stakeholders are fully informed and participate in decision-making related to the public interest in all stages of mining." The Whitehorse Mining Initiative and AQUAMIN recommended the use of Public Liaison Committees as a mechanism for public involvement.

The focus of this section is on the "public" as a stakeholder group, as the public are often not significantly involved in programs such as EEM, but may be able to make important contributions to the EEM program. However, it is important to recognize that the public is just one of several important stakeholder groups with an interest in the metal mining EEM program. For the purposes of this document, a stakeholder is defined as any person or group that has an interest in, is affected by, or has an effect on a watershed in which metal mining occurs, or has a role in decisions made pertaining to that watershed. The public includes all stakeholders with the exception of mining companies or associations, and federal, provincial, territorial, aboriginal or municipal government departments or agencies.

It is important to bear in mind that "the public" is not a homogeneous group, and that within "the public" there may be several different stakeholders, each with different interests and concerns. The public may include, for example: environmental non-government organizations, community groups, commercial and/or sport fishers, concerned individuals, and aboriginal groups.

10.2 Objective and Potential Scope of Public Involvement in EEM

The objective of public involvement in EEM is to ensure that decisions made regarding metal mining EEM are made as a result of informed, inclusive and fair consultation with the public. Effective public involvement in EEM may result in:

- improved EEM study design
- improved decision making in EEM
- increased relevancy of EEM
- increased degree of trust between all stakeholders, and established/improved working relationships
- improved public education resulting in increased awareness and understanding of EEM issues
- improved communications between stakeholders

To achieve this objective and to derive maximum benefit from public involvement, mines are strongly encouraged to facilitate public involvement in a wide range of EEM activities, such as:

Pre-EEM: Before metal mines begin EEM, the public could have input regarding the site-specific objectives of EEM, as well as the site-specific questions to be addressed. This input would help in shaping the overall direction of the EEM program.

Site Characterization: Public involvement during the preparation of the site characterization can be invaluable. Given their knowledge of local conditions, the public may make significant contributions to the description of the study area and possible confounding influences, particularly with respect to the fisheries. In addition, the public may be able to help in the identification of valued ecosystem components (VECs), which are elements of the environment which are valued for biological, scientific, socio-economic, aesthetic, or cultural reasons. VECs may be used to help refine the site-specific objective of EEM, and site-specific questions to asked by EEM. VECs could include, for example, a fish species of cultural or economic significance, a reach of a stream valued for recreational purposes, or a scenic view. If available, such information should be used in designing the EEM study.

Study Design: Following on public involvement in site characterization, public involvement in study design may contribute to the establishment of site-specific environmental quality objectives, the development of reporting procedures, and the identification of suitable exposure and reference areas. In addition, the public may review study designs prior to the commencement of monitoring.

Monitoring Activity: The public may be actively involved in some aspects of the actual monitoring work, particularly sample collection. For example, the U.S. EPA has developed a protocol for the collection of water samples by community groups. In British Columbia, the Pacific Streamkeepers Federation has developed a handbook with includes such topics as water sample collection, stream habitat surveys and stream invertebrate surveys. The existence of such programs point to the fact that, with proper training, the public can be involved in monitoring work. In the field, the public could augment the work of professionals in environmental monitoring. Public involvement can help increase the cost-effectiveness of monitoring, while at the same time increasing the awareness and knowledge of the public. Involving the public in monitoring provides an opportunity for the mine to train and educate the public in EEM, as well as increase awareness and understanding of EEM components.

Data Assessment and Interpretation: The public may review EEM interpretative reports, and have input to decisions regarding the next steps in the monitoring program. This input would be provided in recognition that there are certain aspects of the program, as described the EEM requirements document, for which there is little discretion.

10.3 Mechanisms for Public Involvement

Mines are encouraged to establish mechanisms for public involvement as early as possible in the EEM process, recognizing that mechanisms may evolve as relations with the public change. There is a range of mechanisms by which the public could be involved in metal mining EEM. AQUAMIN and the Whitehorse Mining Initiative strongly recommended the formation of Public Liaison Committees as a mechanism for public involvement. However, the appropriate mechanism for a particular site depends, in part, on the intended scope and degree of public involvement at that site.

A mine may employ more than one mechanism to facilitate public involvement. This may be particularly helpful in cases where a mine wants to use a complementary mechanism to reach a broader segment of the public. It may also be useful in cases where a mine wants to use more than one mechanism to reach different segments of the public that may have strongly divergent interests, or in cases where more than one language is spoken within a community.

It is very important to note that, at sites where effective public involvement mechanisms are already in place, mines are encouraged to use those mechanisms to address EEM issues, rather than establish new mechanisms.

It is essential that a mine commit an appropriate level of financial and staffing resources to support the mechanisms to be implemented. If resources are not adequate, it is unlikely that the mechanisms will be effective. To facilitate public involvement, the mine may have to make some resources available to public participants to cover expenses associated with participation. This may be particularly important in remote areas where there may be costs associated with travel.
Mechanisms by which the public could be engaged include those described below. Note that the mechanisms identified here are not exhaustive. There are other potential mechanisms that a mine may choose to use.

10.3.1 Ongoing Mechanisms

Public Liaison Committee

The Whitehorse Mining Initiative and AQUAMIN recommended that the public is involved as actively as possible, and that sharing authority on certain issues is desirable. Further, AQUAMIN recommended that Environment Canada, in consultation with other stakeholders, develop guidelines for the establishment of public liaison committees, including the reporting of information to the public.

It is strongly recommended that Public Liaison Committees (PLCs) be formed to facilitate public involvement, whenever there is sufficient public interest. Public liaison committees can help ensure that decisions regarding EEM are made in an open, transparent and inclusive manner. If a Public Liaison Committee is to be established, then the following factors should be considered:

1) Stakeholders should be involved as soon as possible and should participate in process design, including:
   - setting the terms of reference for the PLC
   - determining the focus of discussion, within the scope determined by the mine operator
   - identifying participant funding needs, and what expenses may be covered

2) All stakeholders must have a clear understanding of the process, including:
   - the objective of EEM
   - how the public is to be involved
   - the objective of the public involvement process
   - the scope of public involvement, and the scope of decisions to which the public have input
   - who the final decision-makers are, and that the decision will be communicated to the PLC and the general public, complete with reasons for the decision
   - the consequences of not reaching consensus

3) Membership in the PLC should be determined on a site-specific basis. Membership should include representatives of the company and the public. Membership may also include:
   - representatives of the federal government
   - representatives of relevant provincial/territorial/aboriginal government departments or agencies
   - company employees

   Meetings of the PLC should be open to the general public, including anyone wishing either to observe or participate in a specific meeting, thus ensuring that the process remains open and transparent.

4) Accurate, credible, and timely information must be equally available to all participants.

5) Consensus among the participants on the PLC should be the ultimate goal. In the case of consensus not being reached, the PLC would forward its findings, including full discussions of dissenting viewpoints, to the eventual decision-making body. The decision-makers would thus have a clear understanding of the situation and the different options presented, upon which to make a final decision.

Public Liaison Contact

Where there is insufficient public interest to establish a Public Liaison Committee, a mine may identify a member of the public to volunteer to act as a Public Liaison Contact. The Public Liaison Contact should be a person not connected with the mine, either as an employee or contractor, or as a relative of an employee or contractor. It may be difficult to identify an appropriate person to act as Public Liaison Contact, but care should be taken to try to ensure that the person identified is acceptable to the public.

The Public Liaison Contact would receive copies of EEM related correspondence between the mine and the AO, as well as other relevant documentation, including any information prepared by the mine which is specifically intended for the public. The Public Liaison Contact would be the primary point of contact with the public for the mine, and would play a role in ensuring broader distribution of relevant documentation to the public. The Public Liaison Contact could also assist the mine in planning and implementing short-term and/or complementary mechanisms for public involvement.

10.3.2 Other Mechanisms

Open houses: drop-in events designed to allow the public to obtain information and respond at their convenience. Open houses may consist of a visual display, together with handouts and knowledgeable staff to answer questions and solicit opinions.

Public meetings: provide an opportunity to inform the public and for them to make formal and informal presentations, and to exchange comments. To be effective they need to follow an agenda. A representative of the mine, or a neutral party should chair public meetings

Workshops: carefully planned forums designed to air certain issues and to share different points of view. Workshops are usually limited to a small number of invited participants. A facilitator, whose role is to encourage dialogue, structure input towards the workshop goals, and summarize the results, may chair the workshops.

Community visits: mine staff visit community groups to interact directly with local citizens. It provides an opportunity to interact with the public on their "turf" which may be less threatening and more conducive too constructive informal dialogue. This may be a valuable option to consider in remote areas.

Site visits: interested participants visit a mine site to obtain firsthand information and orientation. The visit provides an opportunity for direct contact and exchange of information with the public. The visit also provides the public an opportunity to see the mine site, allowing for a greater understanding of the project.

Electronic communications: internet sites or other means for mine operator to make information available and receive feedback. These provide a cost-effective means of making large amounts of information available, and simplify logistics, since there is no need for building address lists, copying, and mailing out documents. However, appropriate in–house technical expertise is required. Electronic communications should not be used exclusively, since persons without access to the necessary computer hardware and software will not be able to participate.

10.4 Public Involvement Plan

To help in the development and implementation of public involvement processes, mines are encouraged to prepare public involvement plans. The first plan should be prepared as early as possible in the EEM program, and plans should be updated periodically as the EEM program and public involvement activities at a site evolve. Mines are encouraged to provide copies of public involvement plans to the Authorization Officer. Plans provided to the Authorization Officer will be kept on file so that there is a public record of plans submitted.

The objective of the public involvement plan is to outline mechanisms by which the mine proposes to provide information on EEM to the public, seek input from the public, and respond to input from the public.

Note that as part of the public involvement plan, the person with the mine who is responsible for public involvement activities should be identified.

10.4.1 Provision of EEM Information to the Public

The timely provision of information to the public is essential to public involvement processes. As a minimum, mines should make available to the public the executive summaries and text of each public involvement plan, EEM study design, and EEM interpretative report. In addition, mines are encouraged to provide to the public any other documentation that may be helpful.

The public involvement plan should include the following:

a) Description of proposed mechanisms for informing the public that information is available, and for distributing information

The public involvement plan should include descriptions of how the mine proposes to inform the public that information regarding the EEM program exists, and how the mine proposes to distribute this information to the public. Options for informing the public and distributing information include:

- advertisements in local media, including newspapers and radio
- notifications on mine website
- notifications to community organizations, local government, resources users, and labour unions
- information kiosks at community centres or meeting places (e.g. shopping malls, town halls)
- making information available in local libraries, etc.

b) Description of measures that may be taken to provide information in a form that is understandable to the public

In order to help the public understand information such as study designs, interpretative reports and other documents consideration should be given to:

- the language(s) of the community
- education level of the public targeted in the communications
- sensitivity of appropriate cultural communication styles
- acknowledging localized sensitivities in communicating with various segments of the public

10.4.2 Proposed Mechanisms for Public Involvement

The public involvement plan should include the following:

a) Description of current conditions at the site

A description of current conditions at a site with respect to the following factors is important in providing a context for the public involvement plan. Gathering such information may be very helpful to a mine planning public involvement activities, since this will help in establishing realistic expectations for those activities.

- degree of previous or current public involvement activities
- location and size of potentially affected communities
- age, size and history of the mining operation

- nature of historical or current environmental concerns
- human use of aquatic resources in the water shed

b) Proposed objective and goals of public involvement

The overall objective of public involvement for metal mining EEM is to ensure that decisions made regarding metal mining EEM are made as a result of informed, inclusive and fair consultation with the public. Site-specific objectives and goals may be established. Clear objectives and goals will provide a basis for more effective relationships established with the public. Identifying desired results of public involvement will assist not only in the design and implementation of the public involvement plan, but also in the evaluation of the effectiveness of public involvement.

The public involvement plan should also outline the mine's expectations for the proposed public involvement activities, and outline in a clear and transparent manner proposed steps that may be taken in the event that expectations are not being met.

c) Proposed scope of public involvement

Define what EEM activities the public may be involved in and what decisions may be influenced by public input. Timing of the preparation of the public involvement plan must be consistent with the proposed scope, particularly if the mine plans to involve the public in pre-EEM activities or Site Characterization.

d) Define principles for public involvement

There are a number of principles upon which public involvement in EEM should be based. Principles appropriate to a particular site will depend on both the scope and the degree of public involvement.

Open and Transparent: Once a public involvement process has been initiated, all decisions related to the scope of the process should be completed in an open and transparent manner, so that all stakeholders involved, including the public, are aware of the decision making process, and the nature of the decisions. Unless the process is open, fair and equitable, agreement may not be reached and, if reached, may not last.

Purpose Driven: Participants need to clearly understand the objectives and scope of the public involvement process, and see a clear need for their involvement. To be meaningful, public involvement processes should take place while options for decisions are still open. All stakeholders involved, including the public, need a common understanding of the objective, and an acceptance that a public involvement process is an appropriate mechanism to achieve this objective.

Inclusive not Exclusive: All stakeholders, including the public, should be given the opportunity to influence and participate in the process. Mines need to identify all stakeholder groups, including public groups, with a significant interest in the outcome, including those who will be affected by the outcome, those who will be responsible for implementing it, and those who could undermine the outcome if not involved.

Voluntary Participation: Stakeholders participate voluntarily. All stakeholders, including the public, must support the public involvement process, and will need to invest the time necessary to make it work.

Flexibility: Flexibility should be designed into the process. Operating within the framework of the metal mining EEM requirements and guidance, stakeholders should be able to work together to design site-specific public involvement processes. The initial process design may evolve as the stakeholders become more familiar with the issues, the process, and each other. It is necessary to incorporate the feedback of participants in an ongoing evaluation of the process.

Equal Opportunity: All stakeholders, including the public, should have equal access to relevant information, and the opportunity to participate effectively throughout the process. Whenever possible, stakeholders should have the opportunity to choose their own representatives to the process.

Respect for Diverse Interests: Acceptance of the diverse values, interests, and knowledge of the stakeholders involved, including the public, is essential. The mine needs to allow time for other stakeholders to explore and develop common interests despite their different values. Increased understanding fosters trust and openness which assists the stakeholders to move beyond bargaining over positions to explore their underlying interests and needs.

Accountability: The participants, including the public, are accountable both to their constituencies and to the public involvement process that they have agreed to establish.

Cost-efficiency: Public involvement processes must be carried out in a cost-effective manner. At the same time a realistic time frame should be considered to allow participants to effectively liaise, consult and exchange information with their respective constituencies

Implementation: Commitment to implementation of recommendations/decisions of public involvement activities, and feedback to the public regarding implementation are essential parts of any agreement on public involvement. Where recommendations/decisions are not implemented, those involved should be informed, and a rationale should be provided.

e) Proposed timelines

Realistic timelines are essential to public involvement in EEM. The public involvement plan should identify deadlines and milestones for key decisions that may be influenced by public involvement. Timelines should provide optimum opportunities for public involvement, while allowing the mine to meet required deadlines and milestones. Mines should be aware of the scheduling and resources constraints of the public. In scheduling public involvement activities, it may be helpful to identify the schedules and availability of the key public stakeholders first, and then the availability of other public stakeholders, where appropriate. When legislated timelines are tight, and there may not be time for adequate public involvement, this situation should be communicated to the public.

f) Proposed degree of public involvement

Following on the recommendations of AQUAMIN and the Whitehorse Mining Initiative, it is recommended that the public be involved to the fullest extent possible at all mine sites. As the degree of public involvement increases, so do the expectations and need for commitment regarding:

- the level of skill and expertise necessary for all participants
- resource requirements (time and money)
- expectations that consideration be given to input throughout the EEM process
- expectations that input would influence the final decisions made by the mine or governing body

Options for the potential degree of public involvement include the following:

Shared Authority: The public is involved in the decision making process, to a degree and extent agreed to by all stakeholders. Within the terms of such an agreement, the public is an equal partner in the decision making process. The public is formally engaged through the establishment of Public Liaison Committees (PLC). However, under such an arrangement, the government and regulated industry, not the PLCs, must ensure regulatory requirements are met.

Joint Planning: The public is engaged through ongoing consultation in all phases of EEM, from objective setting to review of results. This consultation occurs, in part, through Public Liaison Committees, with a mandate and terms of reference agreed to by all stakeholders. In addition, broader public consultation may also occur through the use of other mechanisms. All stakeholders participating are accountable to the consultation process, and there is obligation that mines and regulatory agencies give due consideration to the results of the process.

Ongoing Public Consultation: Public forums and/or outreach activities to local community organizations are held on an ongoing basis. The objective of these efforts is to provide regular updates (including response to input from previous fora) and to obtain regular input from the public. The scheduling and frequency of these effects are at the discretion of the mine, and there is no formal ongoing relationship with the public, although informal relationships may develop.

Public Consultation: Through public forums and/or outreach activities to local community organizations, public input is sought on various EEM issues, including study design. The public also has an opportunity to review monitoring results. No ongoing working relationships with the public are established.

Information Feedback: Information on the status of the EEM program, including monitoring results, is presented to the public at open forums or through some other means. Forums provide opportunities for the public to comment.

Information: Information on the status of the EEM program, including monitoring results, are available to the public, but there is no formal mechanism for public input or involvement in EEM.

g) Resources for proposed public involvement activities

It is essential that a mine ensure that adequate resources are in place to support the proposed public involvement activities. If adequate resources are not committed, even the best plans may fail. It is a mine's prerogative to keep internal resourcing issues confidential, but a mine should provide assurances to stakeholders, including the public, that it has committed adequate resources to the proposed public involvement process.

The public involvement plan should clarify what kind of financial resources may be available to support the participation of public stakeholders, and if so what types of expenses would be covered, to what degree, and by what type of mechanism.

10.4.3 Proposed Mechanisms to Respond to Public Input

Essential to the public involvement plan is proposing mechanisms to respond to public input. Such proposals are needed to provide assurance to the public that their input will be considered. The broader the proposed scope of public involvement activities, and the greater the degree of proposed involvement, the higher expectations will be on the part of the public that the recommendations/decisions of public involvement activities will be considered. Therefore, proposed mechanisms to respond to public input should be developed to a level of detail appropriate to meet the expectations of the public.

Mechanisms to respond to public input should include mechanisms for decision making based on public involvement, and mechanisms for informing the public of decisions made, including rationale for not accepting recommendations from the public.

11. Information Management

Data will be provided to the federal Department of Environment (DOE) by the mines in electronic format as per section 23 of the MMER, where an electronic format has been provided by DOE.

This section will be updated when an information management tool (e.g., electronic reporting system) for the metal mining EEM program becomes available.

Table of contents

12. Alternative Monitoring Methods *12-1*

12.1 Use of Mescosms as an Alternative Monitoring Method *12-1*
12.1.1 Background of Artificial Stream Development 12-1
12.1.2 Applicability within EEM programs 12-5
12.1.2.1 Pulp and Paper EEM 12-5
12.1.2.2 Metal Mining EEM 12-6
12.1.2.2.1Initial and Periodic Monitoring 12-6
12.1.2.2.2Focussed Monitoring 12-6
12.1.2.2.3Investigation of Cause 12-7
12.1.3 Mesocosm Technology 12-7
12.1.3.1 Suitability of Mesocosm Design 12-7
12.1.4 Description and General Operation of Large Transportable Mesocosm System 12-7
12.1.5 Considerations for Site Selection 12-14
12.1.6 Biological Monitoring Study Designs 12-15
12.1.6.1 Overview 12-15
12.1.7 Study Treatments 12-16
12.1.8 Replication 12-17
12.1.9 Sample Sizes and the Role of Effect Size 12-18
12.1.9.1 No available pre-existing (historical) data 12-18
12.1.9.2 Available pre-existing (historical) data 12-19
12.1.10 Fish Monitoring –Effects Assessment 12-20
12.1.10.1 Study Design 12-20
12.1.10.1.1 Species Selection 12-20
12.1.10.2 Biological Endpoints 12-21
12.1.10.3 Timing and duration of effluent exposure 12-22
12.1.10.4 Fish Methods 12-23
12.1.10.4.1General Set-up 12-23
12.1.10.4.2 Maintenance and Monitoring 12-26
12.1.10.4.3 Sampling and Analysis of Fish 12-26
12.1.10.5 Data Assessment and Interpretation 12-27
12.1.10.5.1 General Requirements 12-27
12.1.10.5.2 Statistical Analysis of Parameters 12-27
12.1.10.5.3 Nested ANOVA Analyses 12-30
12.1.10.5.4 ANCOVA Analyses 12-30
12.1.10.6 Data QA/QC and Analysis 12-31
12.1.11 Benthic Invertebrate Community Monitoring Effects Assessment 12-32
12.1.11.1 Study Design 12-32
12.1.11.1.1 Species Selection 12-32
12.1.11.2 Endpoints 12-32
12.1.11.2.1 Biological Endpoints 12-32
12.1.11.3 Timing and duration of effluent exposure 12-33
12.1.11.4 Benthic Invertebrate Methods 12-34

12.1.11.4.1 General Set-up 12-34
12.1.11.4.2 Maintenance and Monitoring 12-36
12.1.11.5 Sampling and Analysis of Algae and Benthic Invertebrates 12-36
12.1.11.6 Data Assessment and Interpretation 12-37
12.1.12 Supporting Measurements 12-38
12.1.12.1 Water Chemistry Endpoints 12-38
12.1.13 Exposure Endpoints 12-39
12.1.14 Other Considerations 12-40
12.1.15 References 12-42

12.2 Use of Caged Bivalves as Alternative Monitoring Method 12-49
12.2.1 Introduction 12-49
12.2.2 Background 12-49
12.2.3 Approach 12-50
12.2.4 Species Selection 12-51
12.2.4.1 Most Commonly Used Taxa 12-52
12.2.4.2 Marine and Estuarine Bivalves 12-54
12.2.4.3 Freshwater Bivalves 12-54
12.2.4.4 Source 12-55
12.2.4.5 Size and Age 12-56
12.2.4.6 Number of Organisms 12-56
12.2.4.7 Handling and Holding Conditions 12-57
12.2.5 Study Design 12-57
12.2.6 Sampling Design 12-58
12.2.6.1 Area and Station Selection 12-59
12.2.7 Replication 12-59
12.2.7.1 Number of Cage Stations 12-59
12.2.7.2 Replication and Pseudoreplication 12-60
12.2.7.3 Applying Replication To Study Design 12-62
12.2.7.4 Timing and Duration 12-65
12.2.7.5 Modifying or Confounding Factors and Supplementary Measures. 12-65
12.2.7.6 Cage Designs 12-67
12.2.7.7 Mooring Systems 12-68
12.2.8 Variables to be Measured 12-68
12.2.8.1 Survival 12-72
12.2.8.2 Growth 12-72
12.2.9 Methods for Test Initiation, Cage Deployment and Retrieval, Test Termination 12-74
12.2.9.1 Test Initiation 12-74
12.2.9.2 Cage Deployment 12-75
12.2.9.3 Cage Retrieval and Test Termination 12-75
12.2.10 Data Analyses 12-76
12.2.11 Tissue Concentrations 12-77
12.2.11.1 Dioxins and Furans 12-77
12.2.11.2 Mercury 12-77
12.2.11.3 Other Chemicals 12-78
12.2.12 Reporting 12-78

12.2.13 References 12-78

List of Tables

Table 12-1Summary of artificial stream applications for assessing the effects of pulp and paper and mining effluents on aquatic ecosystems as required under Canadian Environmental Effects Monitoring 12-3
Table 12-2 Summary of types of information attained from Canadian artificial stream studies in the last decade (after Dubé et al 2001a)........ 12-4
Table 12-3 Effects on fish........ 12-29
Table 12-4 Biological endpoints to be used for supporting analyses. These analyses are for informational purposes and significant differences between exposure and reference streams are not necessarily used to designate an effect. 12-29
Table 12-5 ANOVA Table for Nested Design, with Two treatments, Four replicate Streams per treatment, and *n* Fish or Composite Samples Within Each Replicate Stream 12-30
Table 12-6 ANOVA Table for Repeated Measure (RM) Design with Two Treatments, Four Replicate Streams per treatment, and t Measurements Times........ 12-39
Table 12-7 Evaluation criteria for EEM monitoring alternatives (after Dubé et al., 2001a). 12-41
Table 12-8 Suggested Taxa for Use in Caged Bivalve Studies for EEMs........ 12-53
Table 12-9 Growth Measurements........ 12-72
Table 12-10 Summary of Statistical Procedures to Analyze Effect Endpoints for Caged Bivalve Studies 12-76

List of Figures

Figure 12-1Picture of mesocosm trailer system used for EEM. 12-9
Figure 12-2 Schematic of mesocosm trailer system (not to scale)(after Dubé 2000). 12-10
Figure 12-3 A) mesocosm stream........ 12-12
Figure 12-4 Fan propeller unit for current velocity. 12-13
Figure 12-5 Fish allocation scenarios where A) adults and juveniles are allocated to the same stream or B) adults and juveniles are allocated to separate streams. For actual implementation, streams are randomly allocated to each treatment. 12-25
Figure 12-6 Example of bivalve mooring design 12-70

12. Alternative Monitoring Methods

At some mine sites the fish and benthic invertebrate community monitoring studies may not be appropriate. The reasons for this are site specific, but the most common reasons are the presence of hazardous conditions (e.g. high velocity) or the presence of confounding factors such as other effluent discharges in the exposure area, that will make it difficult or impossible to isolate any effects attributable to the effluent being monitored.

Where mines cannot design the fish or benthic invertebrate community surveys to resolve difficulties associated with confounding influences, mines will provide a scientific rationale and justification and propose cost effective and technically feasible alternative monitoring methods within the study design. Chapters 4 and 5 describe potential alternative biological monitoring methods.

A number of alternative monitoring methods are recommended in this Chapter . Mines may choose other scientifically defensible methods provided that the results can determine if the effluent is having effects on the fish population (growth, reproduction, condition and survival), fish tissue (mercury levels) or the benthic invertebrate community (benthic invertebrate density, taxa richness, the Simpson's Diversity Index and the Bray-Curtis Index). Recommended alternative monitoring methods to the fish survey are mesocosm studies, caged bivalves, and expanded sublethal toxicity testing. For benthic invertebrate community surveys, the recommended alternative monitoring method is mesocosm studies.

12.1 Use of Mescosms as an Alternative Monitoring Method

12.1.1 Background of Artificial Stream Development

The environmental effects monitoring (EEM) program was developed to aid in assessing the adequacy of both the *Pulp and Paper Effluent Regulations* and the *Metal Mining Effluent Regulations* under the federal *Fisheries Act*. These programs have been designed to achieve national uniformity in monitoring of effects, while taking into consideration site-specific factors. Program requirements and monitoring guidance are given in Environment Canada (1992; 1995a; 1995b; 1997; 1998; 2001). Key components of each EEM program include monitoring of fish, benthic invertebrates and supporting environmental parameters (e.g., effluent and water chemistry, sublethal toxicity testing and sediment chemistry) in reference and exposure areas to assess effluent effects. *This assessment is normally conducted using field survey techniques as they are the most environmentally relevant approach for effects assessment.* However, there may be situations where field surveys as recommended do not provide the required information. These situations could include confounded receiving environments and unsafe sampling conditions. If standard EEM field surveys cannot be conducted, use of scientifically defensible, cost-effective, and technically feasible alternative monitoring methods can be considered.

Artificial stream systems or mesocosms are one of the recommended, scientifically defensible monitoring alternatives for EEM programs. See Section 12.2 for the information on the use of caged bivalves. Artificial streams are recommended as monitoring alternatives because years of research and development have demonstrated that they can produce good quality data on effluent effects that fit within the required regulatory context (Culp and Cash 1995; Culp and Podemski 1996; Culp et al., 1996; Culp and Lowell 1998; Culp et al., 2000a; 2000b; Dubé and MacLatchy 2000; Lowell et al., 2000; Culp et al., 2001; Dubé et al., 2001a; 2001b). Since 1991, the National Water Research Institute (NWRI) of Environment Canada has been developing field-based artificial stream systems for assessing the effects of point source effluents on aquatic ecosystems. Field-based artificial stream studies relevant to EEM applications have been conducted in Canada 14 times in 8 years from 1993 to 2000 (Table 1) (Dubé et al., 2001a). Table 2 summarizes the types of information that have been attained from these studies including the range of receiving environments studied, the types of effluents, specific endpoints for different levels of biological organization, and the types of hypotheses that have been tested.

In the Northern River Basins Study (1991-1996) a large artificial stream system was used as part of an integrated environmental assessment program to determine the effects of nutrients and contaminants from pulp mill effluent (PME) on benthic communities in the Athabasca River, AB (Culp and Podemski 1996; Culp et al. 1996; Podemski and Culp 1996; Culp et al. 2001). Similar studies were conducted in British Columbia from 1991-1998 in the Thompson and Fraser Rivers under the Fraser River Action Plan. In these studies artificial stream systems were used to improve the understanding of the effects of PME on freshwater riverine benthos (Culp and Cash 1995; Lowell et al. 1995; Culp et al. 1996; Dubé and Culp 1996; Lowell et al. 1996; Culp et al., 2000a; Lowell et al. 2000). In the late 1990's, artificial streams were used in three industrial pilot studies to assess the effects of PMEs on small-bodied fish in estuarine (Saint John River, NB) Monitoring (EEM) Programs (Dubé 2000b; Dubé et al., 2001a).

Table 12-1Summary of artificial stream applications for assessing the effects of pulp and paper and mining effluents on aquatic ecosystems as required under Canadian Environmental Effects Monitoring

Year	Program[1]	Research Objective	Location	Artificial Stream System	References
1993-94	NRBS	To determine the effect of both nutrients and contaminants in PME[2] on periphytic algae and benthic invertebrate communities.	Athabasca River, Hinton, AB	Model I: Large non-mobile mesocosm	Podemski (1999) Culp and Podemski (1996) Podemski and Culp (1996) Culp et al. (1996) Culp et al. (2001)
1993-94	FRAP	To assess the effects of increasing concentrations of PME (0.25% -10%) on periphyton and chironomids (Diptera: Orthocladiinae).	Thompson River, Kamloops, BC	Small microcosm system	Dubé (1995) Dubé and Culp (1996) Dubé et al. (1997) Lowell et al. (2000)
1993	FRAP	To determine the food-dependent effects of PME on the mayfly (Ephemeroptera: *Baetis tricaudatus).*	Thompson River, Kamloops, BC	Small microcosm system	Lowell et al. (1995) Lowell et al. (1996) Culp et al. (1996) Lowell et al. (2000)
1994	FRAP	To determine the effects of PME (1% and 3%) on periphyton and benthic invertebrate communities.	Fraser River, Prince George, BC	Model II: Large mobile mesocosm	Culp and Cash (1995) Culp et al. (1996) Culp et al. (2000)
1997	Pulp & Paper EEM	To assess the effects of PME (3%) on a small-bodied fish (*Fundulus heteroclitus*) in a marine environment.	Saint John Harbour, Saint John, NB	Model II: Large mobile mesocosm	Dubé et al. (2001)
1997-98	Pulp & Paper EEM	To determine the effects of final PME, in-mill process streams, and a mill process change on a small-bodied fish (*Fundulus heteroclitus*) in an estuarine environment.	Saint John River, Saint John, NB	Model II: Large mobile mesocosm	Dubé (2000) Dubé and MacLatchy (2000a) Dubé et al. (2001)
1999	TSRI	To determine the effects of primary and secondary treated PME on a small-bodied fish	Miramichi River, Miramichi, NB	Model III: Large mobile mesocosm	Dubé et al. (2001)

		(*Fundulus heteroclitus*) in an estuarine environment.			
1999-2000	TSRI	To determine the cumulative effects of PME and municipal sewage on algae and benthic invertebrate communities.	Saint John River, Edmundston, NB	Medium modular microcosm	Culp, Dubé, Cash, MacLatchy, Glozier (unpublished)
2000	TSRI	To determine the effects of MME[3] on juvenile Atlantic salmon (*Salmo salar*).	Little River, Brunswick Mines, Miramichi, NB	Model III: Large mobile mesocosm	Culp, Dubé, Cash, MacLatchy, Glozier (unpublished)
2000	TSRI	To determine the effects of MME on algae and benthic invertebrate communities.	Little River, Brunswick Mines, Miramichi, NB	Medium modular microcosm	Culp, Dubé, Cash, MacLatchy, Glozier. (unpublished)
2000	TSRI	To determine the effects of MME on a food web including algae, benthic invertebrates and slimy sculpin (*Cottus cognatus*).	Little River, Brunswick Mines, Miramichi, NB	Medium modular microcosm	Culp, Dubé, Cash, MacLatchy, Glozier (unpublished)

[1]NRBS: Northern River Basins Study; FRAP: Fraser River Action Plan; EEM: Canadian Pulp & Paper Environmental Effects Monitoring Program (EC 1992); TSRI: Toxic Substances Research Initiative. [2] PME: Pulp mill effluent. [3] MME: Metal mine effluent.

Table 12-2 Summary of types of information attained from Canadian artificial stream studies in the last decade (after Dubé et al 2001a).

Receiving Environment Type	**Effluent Types[1]**	**Endpoints**					**Hypotheses Tested**
		Primary Producers	**Decomposers**	**Benthic Invertebrates**	**Community Metabolism**	**Fish**	
• Freshwater • Estuarine • Marine	PME • BKM • TMP • Sulphite MME	• Biomass • Community composition • Indicator species • Tissue contaminants	• Decomposition rate • Fungal species composition • Fungal sporulation rate	• Abundance and biomass • Community composition • Emergence rates • Richness, diversity, ratio metrics • Developmental indicators (wingpad spread)	• Gross and net primary production • Microbial respiration • Oxygen isotope signatures	• Growth • Mortality • Condition • Liver size • Gonad size • Reproductive hormones • Tissue metals • MFO[2] induction • Contaminants in bile • "Stress responses" (muscle glycogen)	• Effluent effects • Cause of effect (in-mill waste streams, nutrients vs contaminants) • Process changes • Cumulative effects

[1] PME = Pulp mill effluent; BKM = Bleached kraft mill; TMP = thermomechanical pulp mill; MME = Metal mining effluent. [2] MFO = mixed function oxygenase induction.

The most recent phase of artificial stream development has occurred since 1999 under the Toxic Substances Research Initiatives (Culp, Cash, Dubé, Glozier, MacLatchy, unpublished). In this program, stream systems have been developed to assess the effects of pulp mill and mining effluents on riverine and benthos, fish, and aquatic food webs (i.e., self sustaining multi-trophic level applications including algae + benthos + fish (Dubé et al. 2001a).

The objective of this document is to provide guidance on study design and implementation of mesocosm studies as an EEM monitoring alternative to assess the effects of PME and metal mining effluent (MME) on benthic invertebrates and fish. This Guidance Document is intended to provide information on recommended standards of good scientific practice available to meet the outlined EEM requirements. Please note that this Guidance Document is published in a format and is viewed as a document which can, and will, be updated regularly as new information and research becomes available. To ensure that the most recent updates are included in this document, consult our website at www.ec.gc.ca/eem, or contact Environment Canada's, National EEM Office (phone: 819-997-1535; fax 819-953-0461; e-mail: eem-esee@ec.gc.ca).

12.1.2 Applicability within EEM programs

It should be emphasized that mesocosms are a recommended monitoring alternative but their use should only be considered when field surveys cannot be conducted. Situations where artificial streams may be used include confounded receiving environments and where unsafe sampling conditions exist. Examples of confounded receiving environments include areas with the presence of historical effects, the absence of suitable reference areas for comparison to exposure areas, the presence of other effluents, and changes in relevant habitat types that cannot be factored out in the design of a field survey. There are advantages and disadvantages of using mesocosms and these should be weighed against the advantages and disadvantages of using other monitoring alternatives before final selection of an approach. Recommended alternative monitoring methods to the fish survey are mesocosm studies, caged bivalves, and expanded sublethal toxicity testing. For benthic invertebrate community surveys, mesocosm studies are the recommended alternative monitoring method.

12.1.2.1 Pulp and Paper EEM

Under the current *Pulp and Paper Effluent Regulations*, if field surveys cannot be conducted as outlined in the benthic invertebrate and fish survey decision trees then the use of alternatives is not a *requirement* until their use is included within the legislation (Environment Canada 1998). However, Environment Canada is currently developing proposed amendments to the *Pulp and Paper Effluent Regulations* and to the *Port Alberni Pulp and Paper Effluent Regulations* under the *Fisheries Act*. Included within these proposed amendments is the recommended use of monitoring alternatives for effects assessment for fish and benthic invertebrates, including the use mesocosms. These regulatory amendments are proposed to come into force after the completion of Cycle 3 (April 1, 2004).

Although current EEM requirements are applicable until the regulation completes the amendment process at the completion of Cycle 3, mills that cannot conduct field surveys in Cycle 3 can propose the use of alternatives (e.g., mesocosms, caged bivalves) to their Authorization Officers. Mesocosms have been approved for use at some mills as alternatives in Cycle 2 and are proposed for use at other mills in Cycle 3. Thus, alternatives can be used in Cycle 3 if approved by the Authorization Officer even though their use is not *required* until legislated amendments to the *Pulp and Paper Effluent Regulations* are complete (start of Cycle 4).

12.1.2.2 Metal Mining EEM

12.1.2.2.1 Initial and Periodic Monitoring

Mesocosms can be used for assessing the effects of MME on benthic invertebrates and fish when standard EEM field surveys cannot be conducted (Environment Canada 2001). When effects are seen, it is essential that confirmation of these effects be seen in a subsequent monitoring phase (Periodic Monitoring - confirmation). However, at any time during the program, if an effect is seen, the mine can proceed directly to Focussed Monitoring or Investigation of Cause. If no effects are detected in a monitoring phase, it will also be necessary to confirm the absence of effects in a subsequent phase (Periodic Monitoring - surveillance). During the study design of the subsequent phase(s), it is crucial to evaluate whether the mesocosm design previously used, maximized the potential for detecting effects. Factors that need to be evaluated include species suitability, timing of the study, sample sizes, power of the analyses, and the level of effluent exposure in the mesocosm treatments.

12.1.2.2.2 Focussed Monitoring

When effects have been documented and confirmed through the completion of field surveys or mesocosm studies, it will be necessary to document the magnitude and geographic extent of the effect (Focussed Monitoring). The magnitude and geographic extent of effects can be assessed in mesocosms by measuring biotic responses to an ecologically relevant dilution series, by testing effluent effects on ecologically relevant or sensitive biota, or by testing effluent effects on ecologically relevant habitat types. One of the objectives of the subsequent phases may be to document whether other species demonstrate a similar effect (e.g., to assess the magnitude of the effect). If desired, mesocosm study designs can provide flexibility to address Initial or Periodic Monitoring and Focussed Monitoring within the same study (Dubé et al., 2001a).

12.1.2.2.3 Investigation of Cause

Once the magnitude and geographic extent of the effect on fish is known, and previous monitoring has not determined the cause of the effect, identification of the cause of the effect is determined through Investigation of Cause. Mesocosms can be used for investigation of causal mechanisms associated with an effect. Lowell et al. (1995), Culp et al. (1996), Dubé and Culp (1996) and Culp et al. (2001) demonstrated how dose-response study designs could be used to separate the effects of nutrients from contaminants in PME on benthic food webs. In addition, Dubé (2000), Dubé and MacLatchy (2000; 2001), and Dubé et al. (2001b) showed how artificial streams can be used to isolate and identify the contribution of different in-mill process streams to final effluent quality and fish effects.

12.1.3 Mesocosm Technology

12.1.3.1 Suitability of Mesocosm Design

To maintain data quality and ensure consistent application of mesocosm technologies in EEM programs, review of the physical design of the system used, and the study design, including standard operating procedures, is required. This Guidance Document outlines licensing requirements, physical design guidance, and study design recommendations specific to a large transportable mesocosm system developed by NWRI.

Development of the large mesocosm system is on-going. A smaller modular design is also under development for use in EEM programs. As the technologies develop, design improvements and licensing opportunities will be described in amended Guidance Documents.

12.1.4 Description and General Operation of Large Transportable Mesocosm System

The mesocosm system recommended for use was developed by the NWRI and has been used to assess PME and MME effects on benthic food webs (benthic invertebrates and algae) (Culp and Podemski, 1996; Culp et al., 1996, Culp et al., 2000a; Culp et al., 2001; Dubé et al., 2001a) and small-bodied fish (Dubé 2000; Dubé and MacLatchy 2000; Dubé et al., 2001a; 2001b). The stream system consists of 16 circular, 0.9 m^2 surface area tanks or streams placed in pairs on tables that are 74 cm high (Figures 1 and 2). Water from a reference area is pumped into a 378 L polyethylene head tank placed on a 1.2 m high platform, and gravity-fed through a system of pipes to the streams. Intake pumps such as a land-based pump (e.g., commercially available irrigation or pool pumps) or a submersible pump (0.5 HP Hydro-MaticJ Model # SPD50-H) can be used depending upon volume requirements, pumping distance, and the head required. Pumping reference water from a mill or mine water intake is also a possibility. Gate valves control water flow to individual streams, and allow flow rate calibration for each stream. Water is delivered to each stream at a rate of 2 L/min resulting in a total water requirement of 32 L/min

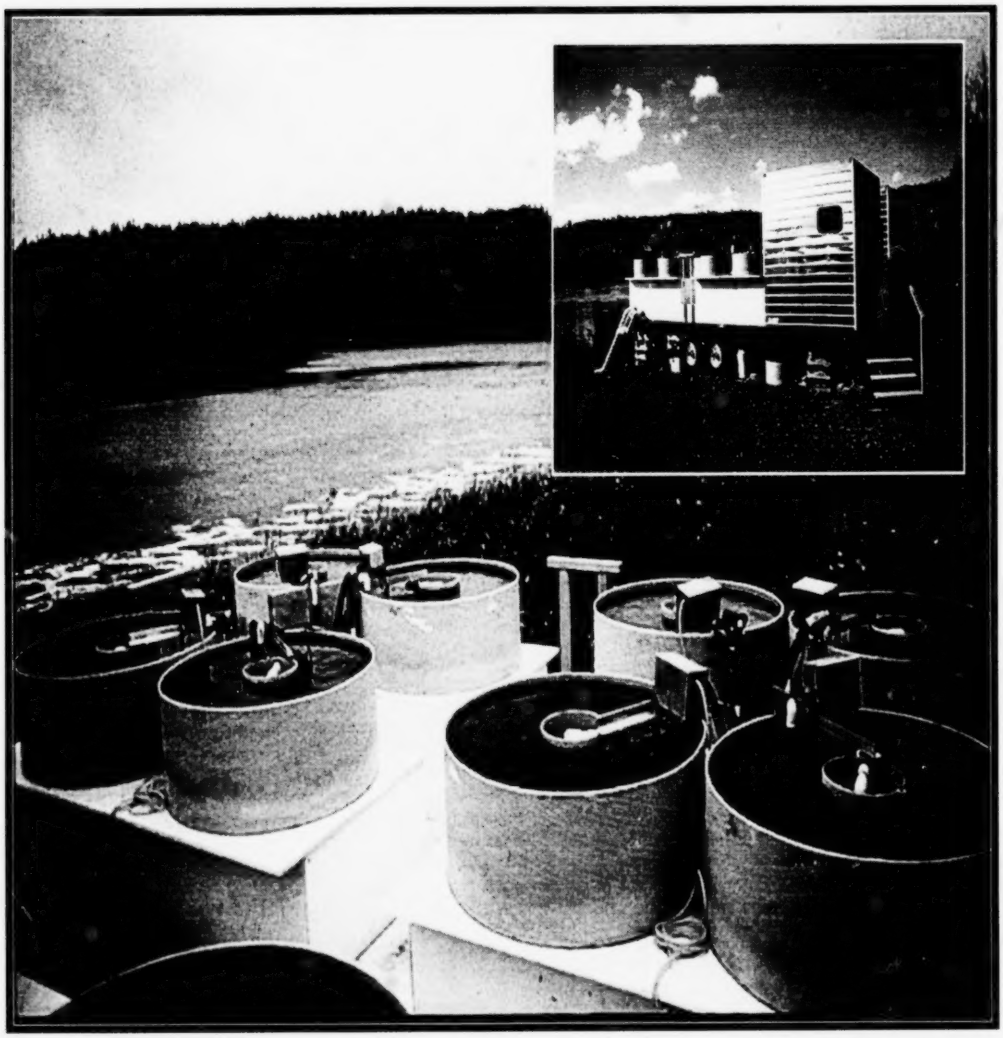


Figure 12-1Picture of mesocosm trailer system used for EEM.

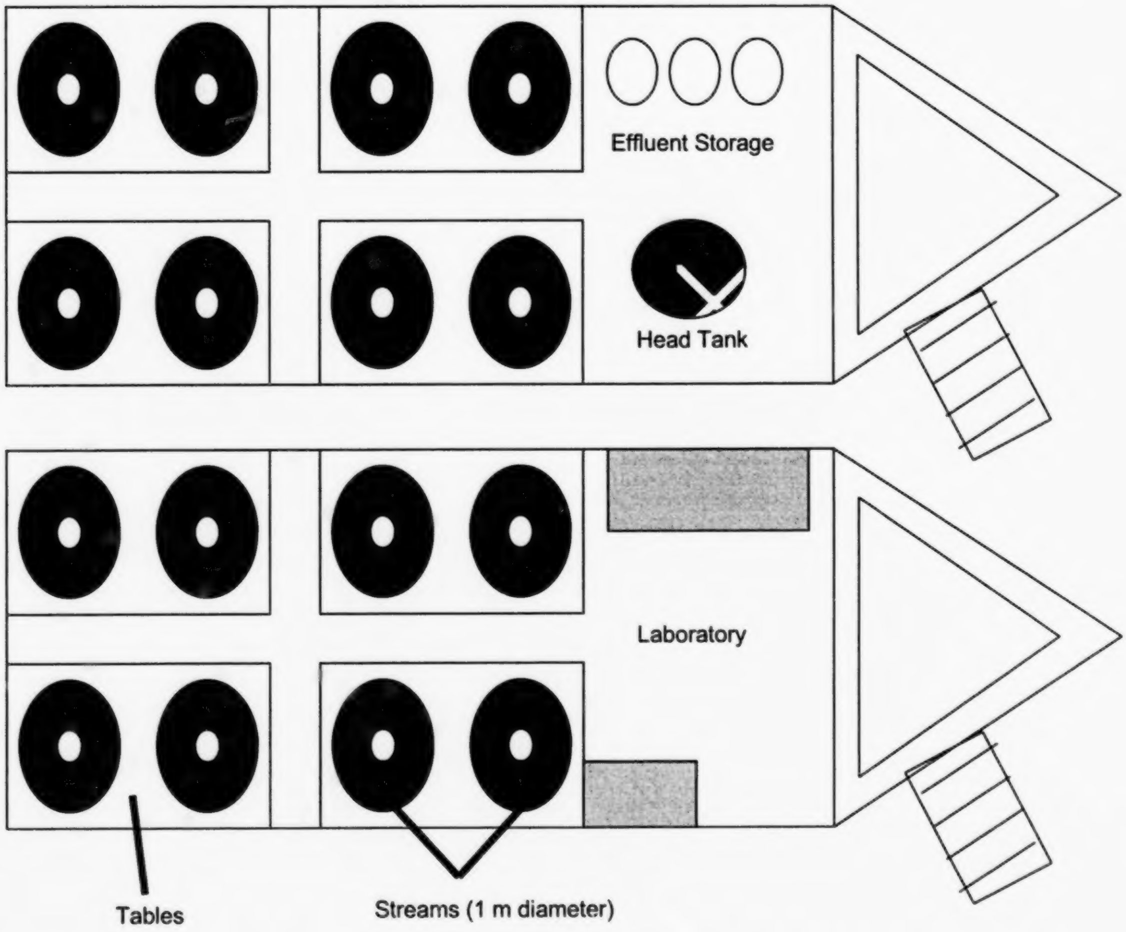

Effluent Storage
Head Tank
Laboratory
Tables
Streams (1 m diameter)

Figure 12-2 Schematic of mesocosm trailer system (not to scale)(after Dubé 2000)

for the 16 stream system (Figure 3a). Water depth in the streams is maintained at 26.9 ± 0.1 cm ($\bar{x}$ ± 1 SE) by an overflow drain that returns all waste water to the river (Figure 3b). The overflow drains are screened to prevent fish loss and limit emigration of insects from the streams (Figure 3c). Each stream contains 227 L of water resulting in a hydraulic residence time of approximately 2 hours. By increasing water residence time within the streams (4 hours), the volume of effluents and reference water required during a study is minimized. However, we recommend an optimal stream residence time of ≤2 h. These recommended residence times are based on previous studies. However, final determination of residence time for a particular system depends upon the size of the tank used and species requirements for oxygen and temperature. If mesocosm studies are conducted in the late autumn, the head tank and water delivery lines can be wrapped with heat tape and insulated to allow the system to be operated when freezing temperatures may occur (-5°C). Shade cloth can be used to reduce solar heating in the summer.

The streams on the large mesocosm system are 107 cm diameter tanks constructed of polyester fiberglass (Figure 1). Streams are constructed by cutting 38 cm long sections of 107 cm pipe and bonding a flat sheet of fiberglass to one end of the pipe. A 25 cm diameter section of pipe is then centered in the larger pipe and the bottom cut out to form a standpipe. Streams are placed on eight, 74 cm high tables, two to a table. The water outflow pipe passes through the standpipe and drains into pipes beneath. These drain pipes connect to a general outflow pipe for discharge to the river downstream of the water intake point.

For invertebrate applications, current velocity in each stream is created by a belt-driven propeller system because this type of propulsive device is relatively gentle on drifting organisms (Figure 4) (Podemski 1999). The motor assembly for each stream is comprised of a geared head motor (250 RPM, 1/40 amp) driving a 22.6 mm pulley. A belt-drive transmits power from the motor to a 49.3 mm pulley mounted on the propeller shaft. The motor and associated electronics are mounted on an aluminum frame in a weatherproof enclosure. This frame is clamped to the top edge (outside) of the stream. A 16 x 230 mm long copper strut extending downward from the aluminum frame holds the propeller shaft bushing and grease seal. A grease nipple at the top of this tube allows for lubrication of the propeller shaft and bushing. The propeller (one per stream) is a 23 cm (9 inch) diameter aluminum fan blade which rotates at a no load speed of 115 RPM. Water velocity in each stream near the water column midpoint is normally maintained at 20 cm/sec although site-specific velocities should be the determining factor. Other current generation mechanisms can be used providing they produce the water velocities. Current for fish studies can also be attained using the propeller system, although spray bars attached to the water delivery system for each stream have also been used for species where water velocity is not a critical requirement.

Effluent is collected daily or every second day and stored in polyethylene containers. The point of effluent collection depends upon the study design. For an Initial Monitoring study at a pulp mill for example, the final effluent that is representative of what is being discharged to the aquatic environment is the target source. Effluent treatments are delivered independently and continuously to individual streams by peristaltic pumps

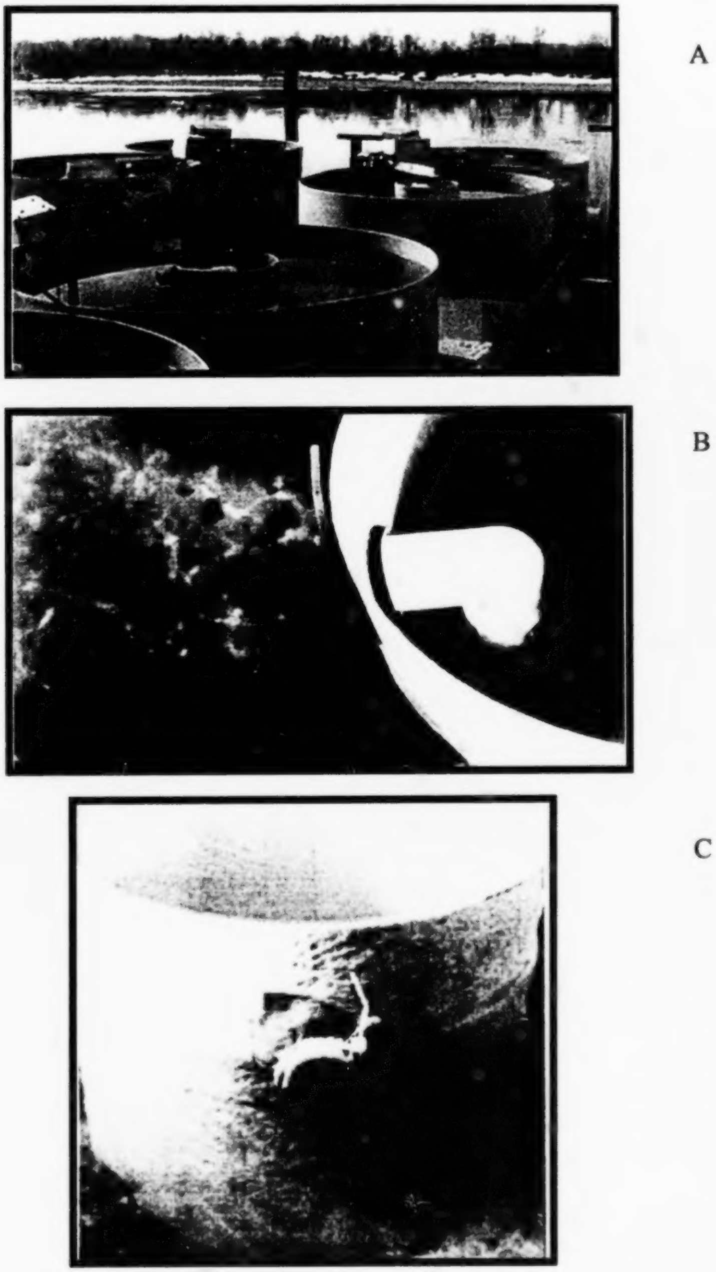


Figure 12-3 **A) mesocosm stream**
B) stream overflow drain on a mesocosm stream
C) screened overflow drain on a mesocosm stream

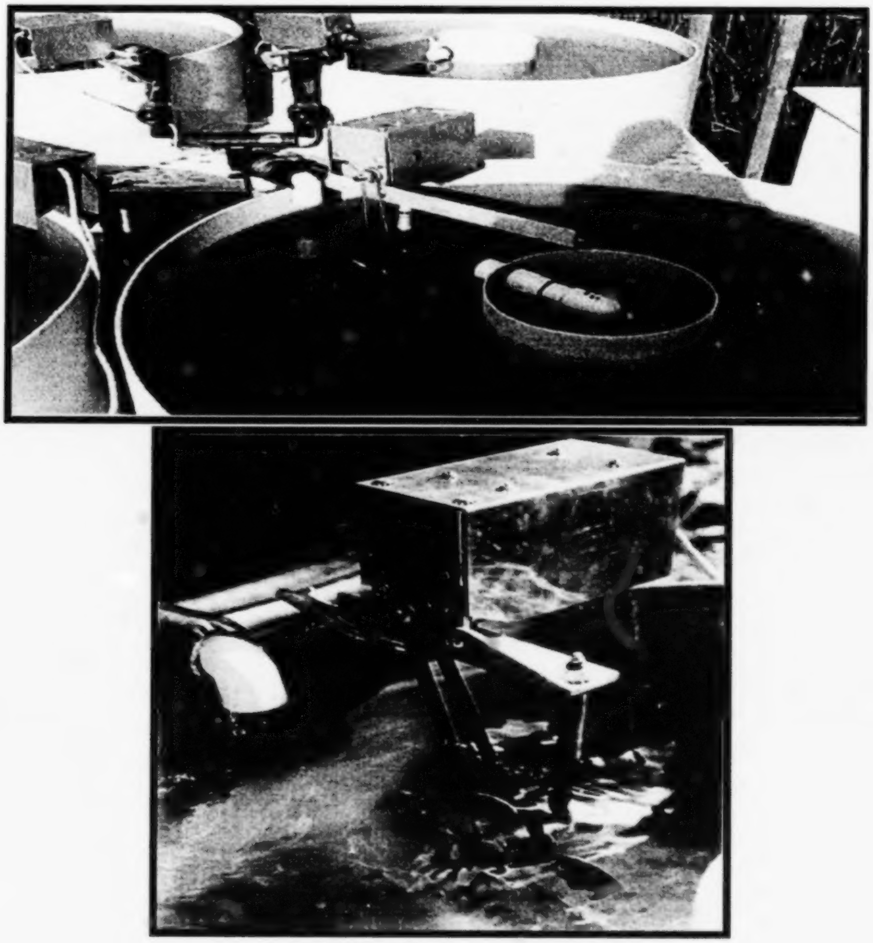


Figure 12-4 Fan propeller unit for current velocity.

(Masterflex ® L/S Nema-type 13 wash down controllers and cartridge pump heads) and a series of insulated tubes for solution delivery. Peristaltic pumps can be kept in insulated boxes to keep them within approved operating temperatures. Tubes carrying effluent can be threaded through foam pipe insulation to the streams, then fed into the water delivery spout on the water delivery pipe for each stream. Effluent flow rates depend upon the site-specific, environmentally relevant concentrations to be tested. For example, if plume delineation studies in the field have determined effluent concentrations to be 1%, then an effluent flow rate into each stream is set at 20 ml/min (1% of 2 L/min). If there are 16 tanks on the trailer, which are allocated between two treatments (8 streams for control; 8 streams for 1% exposure), then the following effluent volume is required on a daily basis: 20 ml/min X 60 min/hr X 24 hr/d = ~29 L/d/stream X 8 streams = ~231 L/d. This volume would fit into a small polyethylene container.

After water and effluent flows have been calibrated for each system, the individual tanks or streams are seeded with natural substrates, algae, benthic invertebrates, and/or fish species endemic to the receiving environment being studied. Inoculation of biotic populations is described below for fish and invertebrates. Further details on construction for this trailer system can be attained by contacting Dr. Joseph Culp (joseph.culp@ec.gc.ca) or Dr. Monique Dubé (monique.dube@ec.gc.ca) at the National Water Research Institute, Environment Canada, 11 Innovation Blvd., Saskatoon, SK, S7N 3H5.

12.1.5 Considerations for Site Selection

In order to operate the field-based mesocosm system there are some basic site requirements. The system is typically located at a reference area for access to reference water. In a freshwater riverine situation, reference water is pumped from upstream of the effluent discharge into the system and discharged back to the river downstream from the point of intake. This is the most straightforward scenario for mesocosm use. If the mesocosm is used in marine or estuarine applications, the site selected for the study must ensure that the quality of the reference water supply is not contaminated by effluent due to tidal influences. Effluent contamination includes both effluent discharged from the mill or mine, or effluent discharged from the mesocosm system itself (i.e., outflow from the streams). In the former case, plume delineation studies are useful to determine the extent of effluent delineation in the receiving environment for appropriate selection of a reference site outside of the zone of effluent influence (Refer to Environment Canada 1998; 2001). In the latter case, the mesocosm can be located on a breakwater, land spit, or wharf to ensure the mesocosm outflow does not contaminate the inflow.

In mining applications, often the mine site is located in river headwaters and adequate reference areas can be a substantial distance from the mine site, particularly in remote areas. In addition, mining effluent concentrations are typically much higher than that found for pulp and paper discharge scenarios resulting in effluent requirements exceeding reference water requirements. Under these conditions, locating the mesocosm system at the mine site and hauling and storing reference water is possible. In this scenario, the plumbing infrastructure on the trailer can be used for effluent delivery and the peristaltic pumps for reference water delivery.

Distance to the reference water source is also a consideration for site selection as extensive pumping requirements with respect to distance, or elevation (head) can exceed pump specifications for the rate required. Other site requirements include adequate space, site access, power and security. A 9 x 5 m area of level ground with good drainage is required for the trailer system. Access to power is one of the major site selection requirements. The system described in this document requires 120Amps of 115 Volt to operate the intake pump, the current generators, air pumps, etc. The mill or mine typically provides this by installation of a power panel as an extension from an existing power line. In remote areas, if power is not available, generators have been used to power the system although this approach is not recommended. This is because maintenance and supervision of generators can make this approach cost prohibitive.

Once the site has been selected for the mesocosm study, the trailer is transported to the site, and water and power are hooked up. It is important that the trailer be properly oriented to minimize stream differences associated with light and wind exposure. One of the key benefits of mesocosms is that biota are exposed to effluents under environmentally relevant conditions of ambient temperature, water quality, and light regimes. If possible, it is recommended that the trailer be located in a north-south direction perpendicular to sunset/sunrise. The trailer should also be located away from the influences of shadows from buildings, etc that are in close proximity to the study area (i.e., all streams should receive the same duration and quality of sunlight).

12.1.6 Biological Monitoring Study Designs

12.1.6.1 Overview

The objective of a study design is to outline what mesocosm and associated laboratory work is needed to complete the biological monitoring portion of the EEM study. Many of the study components are similar to those outlined in the Technical Guidance documents for field surveys (Environment Canada 1998; 2001).

Prior to any biological monitoring studies being conducted, a study design will be submitted to the Authorization Officer at least six months prior to the commencement of conducting the biological monitoring program. The study design will include:

- a detailed timetable for conducting the mesocosm study;
- a description of the study area;
- a description of how the fish studies will be conducted in order to determine if there are effects on the fish population and fish tissue;

- a description of how the benthic invertebrate community studies will be conducted in order to determine if there are effects on the benthic invertebrate community as an indicator of effects on fish habitat; and,
- identification of the quality assurance and quality control measures that will be taken to ensure the validity of the data.

Other recommended details of the study design may include:

- defining the goals and expectations of the EEM study;
- determining the overall approach, including stating the rationale for choosing an alternative, which may be based on previous monitoring results;
- establishing statistical design criteria: development of hypotheses, selection of statistical methods, determination of data needs (statistical significance and power analysis);
- describing the design of the monitoring study: where to sample, what to measure, number of samples;
- developing operating plans and procedures: sampling procedures, laboratory analysis procedures, quality assurance and quality control procedures, data storage and retrieval, data analysis; and
- describing a plan for data interpretation and program evaluation.

The concepts and basic elements of a good measurement system, quality assurance/ quality control (QA/QC) are discussed in Environment Canada (1998; 2001). QA/QC procedures should be described in the study design and prior to the study. Quality assurance results that may affect the reliability of the conclusions will be submitted as part of each interpretative report.

12.1.7 Study Treatments

In mesocosm studies, the effect of effluent on fish is evaluated by comparing population indicators (growth, reproduction, condition and survival) between fish held under control conditions (reference water) to those held in effluent. For metal mining EEM programs, the effect may also be assessed by comparing levels of mercury in fish tissue between reference and exposure species. In mesocosm studies using benthic invertebrates, an effect is determined by comparing measurements of benthic invertebrate density, taxa richness, the Simpson's Diversity Index and the Bray-Curtis Index between invertebrates in reference versus exposure streams. It is crucial that mesocosm studies be designed to maximize the possibility of detecting effects if they are present. This includes using a *mesocosm design licensed for use by Environment Canada*; and selecting appropriate treatments, level of replication, sentinel species, endpoints, and conducting the studies at the proper time of year.

Several study designs can be employed depending upon site-specific requirements, and the phase of the monitoring program [i.e., Initial Monitoring, Focussed Monitoring, and

Investigation of Cause]. The multiple tank mesocosm design provides flexibility for addressing site-specific considerations in the EEM study design with the 16 tank system resulting in up to three different treatments that could be assessed at the same time (five replicate tanks per treatment) under environmentally realistic exposure conditions. A control/impact design can be used where biota (i.e., benthic invertebrates or fish) held under reference conditions are simply compared to biota exposed to an environmentally relevant concentration of effluent (i.e., reference vs exposure). It is recommended that the environmentally relevant concentration of effluent is based upon a plume delineation study conducted during site characterization and represents the effluent concentration in the near field area after complete mixing (Environment Canada 1998; 2001). In the control-impact design for a fish program, typically adult (20 sexually mature males and 20 sexually mature females) and juvenile (20 sexually immature) fish of one small-bodied sentinel species are allocated to each stream (eight streams per treatment). Benthic invertebrates are stocked in the system at ambient densities found at the reference area in the dominant habitat (Section 3.4).

In some cases, additional treatments may be desired. Dose-response study designs where an additional and higher effluent concentrations are used, is helpful to confirm biota responsiveness and the absence of effects. For example, Dubé and MacLatchy (2000) used three treatments (control vs 1% vs 5%) to assess the effects of a final effluent discharged from a bleached kraft pulp mill on a saltwater killifish (mummichog: *Fundulus heteroclitus*). In this example, effects were not observed at the environmentally relevant effluent concentration (1%) but did occur at 5%. This design proved invaluable to demonstrate that the fish were capable of responding to the effluent and the absence of effects at the 1% concentration was real and not a function of short-duration exposure studies. Similarly, Culp et al. (2000) used three treatments (control, 1%, 5%) to assess the effects of treated bleach kraft effluent on benthic food webs.

Additional treatments may also be desired if more than one mine discharges into the same sampling area. Each discharge can represent an exposure treatment in the experimental design and the effects of each effluent can be examined in isolation or in combination. For example, if a mine discharges three effluents into the same receiver then effluent effects can be evaluated using the following treatments: control, effluent 1 (environmentally relevant concentration), effluent 2, effluent 3, effluent 1+2+3 (environmentally relevant concentration). In this example, the number of streams used per treatment would decrease to three.

12.1.8 Replication

There has been considerable discussion surrounding the level of replication appropriate for mesocosm designs. For benthic invertebrate studies, individual streams are the unit of replication for the community-level endpoints. Thus in a control/impact design employing two treatments (reference vs exposure), n=8 as there are 16 tanks on the mesocosm trailer. If additional treatments are preferred (see below), then 5 replicate streams per treatment should be the minimum number of replicates for EEM studies as this should provide adequate power for assessing effects when streams are the level of replication and when no previous monitoring data is available. This approach is consistent with the recommended method for determining the number of sampling stations, using statistical power in field survey designs (Environment Canada 1998; 2001).

The unit of replication is less clear when the system is used to measure individual-based endpoints in fish. The basis of the decision lies in quantification of the importance of a tank or stream effect (*i.e.*, an effect of one tank relative to another within the same treatment). In many laboratory studies where fish are held in aquaria, an assumption is made that there is no biological reason for tank differences within a treatment and thus all individual fish measurements are pooled within a treatment (*i.e.*, the variability attributable to tank effect is often ignored). For mesocosm studies we suggest that a tank effect may exist and the variation attributed to that effect requires some consideration in the statistical design. This is especially the case for longer-term exposures (e.g., 60 days). A nested ANOVA could be used in this example where the tank effect is nested within the fixed-effect factor of effluent treatment. If a tank effect exists then the level of replication should be at the stream level. If tank effects do not exist then fish can be selected as the unit of replication and pooled across streams within a treatment resulting in significant increases in the number of replicates. For regression-based analysis (ANCOVAs) any tank differences would likely fall out as outliers in the analysis. A description of the recommended statistical design for fish is outlined in Section 12.1.10.5.2.

12.1.9 Sample Sizes and the Role of Effect Size

12.1.9.1 No available pre-existing (historical) data

In mesocosm studies using benthic invertebrates, sample size represents the number of streams replicated per treatment as streams are the unit of replication. In a control-impact design, 8 streams are replicated per treatment. If additional treatments are required, then a minimum of five replicates or samples are required for each treatment to address power requirements without pre-existing data (see below). Each sample represents a mean of subsamples collected from each stream.

For a standard fish population field survey, where there is no background monitoring data, the minimum sample size recommended is 20 sexually mature males and 20 sexually mature females of two fish species collected from each of the reference and exposure areas. If small-bodied fish species are chosen as one or both of the fish species, an additional 20 sexually immature fish should also be collected. The rationale for using 20 fish of each sex is that there is little change in the 95% confidence limits with increasing sample size beyond 20 fish.

The current approach for mesocosm studies is to allocate 15 - 20 sexually mature males and 15-20 sexually mature females of one small-bodied sentinel species to each tank. In addition, 20-30 juvenile fish are allocated to the same tank. These are recommended sample sizes that can be increased or decreased depending upon the species selected, the statistical power of the study design, and fish variability. If tanks within a treatment are selected as the unit of replication then 20 fish per sex per tank provides a mean with good precision (i.e., there is little change in the 95% confidence limits with increasing sample size beyond 20 fish). If tank effects do not exist and error terms are pooled so that fish become the unit of replication, then 15-20 fish per sex per tank provides much higher replication per treatment. For example, a study design of control vs 1% effluent with 8 streams per treatment, and 20 males, 20 females, and 20 juveniles per tank can result in 160 males, 160 females, and 160 juveniles per treatment if fish are the unit of replication.

12.1.9.2 Available pre-existing (historical) data

If data are available, then the sample sizes to measure a certain effect size in a parameter with a targeted level of statistical power can be calculated because sample variability is known. Larger numbers of samples are needed if parameters are highly variable, if detection of small differences between reference and exposure streams (small effect size) is desired, or if a high level of power is required. To determine the sample sizes using pre-existing data, the effect size and the statistical power level that is acceptable for the decision-making process needs to be determined. The purpose of defining an effect size and power level is for determining when the sampling program is collecting sufficient information for decisions to be made. Sample sizes can be calculated using methods described in Environment Canada (2001). Appendix 1 of Environment Canada (1997) provides a detailed discussion of power relationships, effect sizes and the benefits of reducing variability in terms of increasing power. It is recommended that α and β be set equal at 0.10 as outlined in EC (1998; 2001) for power calculations.

The greatest factor affecting statistical power is sample variability. In general, mesocosm studies provide less sample variability compared to field studies as exposure is controlled and treatments are replicated. Mesocosm studies are designed site-specifically and priority should be given to reducing variability rather than increasing sample sizes. Variability can be reduced by randomly allocating streams to treatments, and biota (both benthic invertebrates and fish) to streams as described below in Sections 12.1.10.4.1 and 12.1.11.4.1. Random collection of samples at the end of the exposure period also helps to reduce variability to avoid collecting all samples from the same treatment at the same time (e.g., all control streams in Day 1 and exposure streams in Day 2). Other efforts to decrease variability include appropriate timing for the study, species selection, and QA/QC in sampling and analysis protocols. In addition, for fish studies, fish of uniform size class are typically allocated to each stream to reduce variability. However, the use of small-bodied fish species, especially fractional spawners, may increase variability due to differences in spawning condition amongst individuals at the time of the study.

Composite tissue samples for chemical analysis can also increase sample variability if the number of composites per stream is low.

12.1.10 Fish Monitoring –Effects Assessment

12.1.10.1 Study Design

12.1.10.1.1 Species Selection

The most important factors when selecting fish species for mesocosm studies are environmental relevance, abundance for collection, size, availability of adults and juveniles, sexual dimorphism, spawning period, and sensitivity to effluent. Local input may be valuable in providing information on the presence of fish species or information that may be helpful in determining sample locations for fish collection. The recommended method to assess effluent effects in mesocosm studies is by monitoring pre-spawning adults (sexually mature fish) and juveniles of one small-bodied fish species (e.g., darters, minnows, sculpins) that is relevant to the receiving environment. Monitoring of adults provides for assessment of effluent effects on survival, energy use (i.e., reproductive endpoints) and energy storage (i.e., condition, liver size). Monitoring of juveniles provides for assessment of effluent effects on survival, energy use (i.e., growth as a size-at-age substitute) and energy storage (i.e., liver size and condition).

The advantages and disadvantages of using small-bodied species in field surveys has been described by (Gibbons et al., 1998a; 1998b). A small bodied fish can be considered as a fish species that has a maximum size of 150 mm or less. Information on maximum growth of fish species can be found in the scientific literature, including Scott 1967; Scott and Crossman 1973; Fritz et al., 1975; Roberts 1988; Nelson and Paetz 1992; Jenkins and Burkhead 1993; Coad 1995; Leblanc and Couillard 1995; Leblanc et al., 1998. Shuter (1990) noted that small bodied fish species were more sensitive to acidification than larger species or at least responded more quickly to changes in pH. On a practical level, small bodied fish species are usually more abundant, easier to capture, and more sedentary than larger bodied fish species. In addition, measurements for the study can be made precisely and accurately on these small-bodied species.

There can be disadvantages to using small bodied fish. Often, less is known about their basic biology, particularly their spawning habits, making it difficult to determine the best sample areas, times and methods. Some species are also multiple spawners (produce more than one clutch of mature ova every year) (e.g., Heins and Rabito, 1986; Burt et al., 1988; Paine, 1990). This can be a disadvantage for field surveys because reproductive effort in these species is difficult to estimate from a single sample. Reproductive tissue can be turned over almost completely between clutches (i.e., most of the mass of ova in the ovary will be spawned, and then a new clutch of mature ova will be developed). The number of clutches produced during the spawning season becomes the important reproductive variable and is difficult to estimate for an individual female in the field, even with frequent sampling. However, multiple spawners such as the mummichog (*Fundulus heteroclitus*) have been used successfully in mesocosm studies to evaluate effluent effects on EEM endpoints including changes in gonad size (Dubé and MacLatchy 2000; Dubé et al., 2001a; 2001b). An advantage of using controlled exposure mesocosm studies is that the state of fractional spawners can be monitored throughout the exposure period.

In addition to monitoring the effects of effluents on a single small-bodied fish species, the flexibility of the mesocosm system also allows for monitoring of two small-bodied fish species, or juveniles of large finfish. Theoretically, large-bodied finfish could theoretically also be used in mesocosm studies, although this approach has not been used previously. Species selection for mesocosm studies is restricted by size requirements (numbers of fish per tank decrease with increasing fish size), the number of treatments (increased numbers of treatments decreases the number of tanks available per treatment and limits the number of fish that can be monitored), and sexual dimorphism (if males and females cannot be externally sexed, then increased numbers should be considered to ensure adequate sample sizes). Sampled species should also be suitable for measuring the recommended variables. To date fish species used in mesocosm studies include mummichog (*Fundulus heteroclitus*), juvenile Atlantic salmon (*Salmo salar*), and slimy sculpin (*Cottus cognatus*). Fish proposed for use in 2001 and 2002 studies include creek chub (*Semotilus atromaculatus*), rainbow trout (*Salmo gairdneri*), longnose dace (*Rhinichthys cataractae*), and fathead minnow (*Pimephales promelas*).

12.1.10.2 Biological Endpoints

The endpoints measured in a mesocosm study using fish are the same as those measured in a field survey. To answer the question, "has the fish population been modified by effluent?", effects on growth, reproduction, condition and survival of the fish population are determined. Survival, energy use (growth and reproduction) and energy storage (condition, liver size) are measurements that can detect an effect in the fish population related to effluent. All of these measurements can be taken from fish exposed to effluents in mesocosm studies.

Growth is the change in size (weight or length) with time or age. Mesocosm studies over a longer exposure period (45-60 d) can assess effluent effects on juvenile growth, specifically changes in total body weight and length relative to control fish. In mesocosms, growth is measured in Y-O-Y or juvenile fish from the start of the study to the end; a duration of 6 to 8 weeks depending upon the temperature during the study. This is a more direct measure of growth effects than size-at-age estimates used in field surveys. Determination of growth effects over a short exposure period requires that studies be conducted during warm water temperatures and with a sentinel species that has high growth rates. Kneib and Stiven (1978) have shown that juvenile growth rates of mummichog for example, are very high post-hatch (an increase of 15 mm in two months) which make them an ideal candidate for assessment of the effects of effluents on growth. More recent studies using Y-O-Y slimy sculpin have shown growth effects can be measured after only 26 days of effluent exposure (Culp, Cash, Dubé, Glozier, MacLatchy, unpublished). Growth measurements in juveniles can be improved further by tagging individual fish (fin clips, ink dots, branding) for length and weight measurements. Tracking the growth rate of specific individuals decreases the variability of these measures increasing the ability to assess effects. Growth can also me measured throughout the study at regular intervals if desired. Although length and weight measurements are taken in adults, growth is unlikely during the short duration of the mesocosm studies.

Reproduction is expressed as reproductive effort, fecundity, egg size or gonad weight relative to body size. To date mesocosm studies have examined changes in gonad size to assess effluent effects on reproductive function. However, measures of fecundity and egg size are easy to measure, if an appropriate sampling time is chosen. Ideally exposure studies should commence 6 to 8 weeks prior to the spawning season to assess effluent effects on gonad weight.

During the mesocosm study, the physical state of the fish is also assessed. Condition can be used as an index of well being and should be reported. A visual estimation of physical malformations including lesions on the body surface including: eroded, frayed or hemorrhagic fins; parasites; or other physical deformations, is recommended.

12.1.10.3 Timing and duration of effluent exposure

Timing of the mesocosm study is important and dependent upon the spawning cycle of the fish species. Ideally, for spring spawners, studies will be conducted 6 to 8 weeks before spawning commences. For early spring spawners for which it is impossible to study for this length of time prior to the spawning season, studies should be collected as late in the year as possible to allow for gonadal senescence and recrudescence. For fall spawners, a spring or summer survey is appropriate. This may not apply to fish in which ova mature rapidly. For example, some late-spring-spawning minnows should be studied in early spring, rather than in fall, when ova may still be immature. Water temperatures are also a consideration for timing of studies where juvenile growth rates are higher during the summer months. If historical data exists, it would be useful to examine the data and, if appropriate, conduct the study during similar periods so that the studies can be compared. Similarly, subsequent monitoring should be conducted during the same periods of the year to be comparable.

The duration of the mesocosm studies is a balance between being long enough to measure changes in the required endpoints and short enough to be cost-effective. Exposures of 30 d during summer and fall months have been used to measure changes in adult organ size (liver size, gonad size) and growth rates of Y-O-Y slimy sculpin as a result of effluent exposure. Longer exposures 50-60 days are recommended for assessment of both adult endpoints as well as changes in juvenile growth. Exposures beyond 60 days are not considered cost-effective unless cost reduction measures are taken (e.g., hiring a local summer student to maintain the system during the study).

All mesocosm studies should be conducted during normal industrial operations. Effluents are collected daily or every second day during the studies and should be representative of normal operating conditions. Storing of effluents during hardwood to softwood furnish shifts or during periods of bleach plant shut downs for example, should be avoided. Studies should be conducted when the mill is using a consistent furnish (i.e., half the study shouldn't be conducted during a softwood run and the other half during a hardwood run). Mesocosm studies at mines should not be conducted when effluent has not been discharged for long periods (months) or during waste water treatment upsets.

12.1.10.4 Fish Methods

12.1.10.4.1 General Set-up

Once the mesocosm system has been transported to the study site and reference water and power hooked up, then substrate is added to the streams, and water flow rates and stream current velocity are calibrated as previously described. Depending upon the site specific substrate characteristics, washed, crushed gravel or sand can be placed into each tank to a depth of 5 cm. In previous studies, large washed rocks were also added to each tank to serve as refugia for the fish.

Fish are collected from a reference area that is not exposed to the effluent being studied. Obviously, collection requires non-lethal sampling techniques including minnow traps, trap nets or electrofishing with barrier nets. These techniques are described in detail in the Technical Guidance Documents (Environment Canada 1995b; 1998; 2001). Fish are usually sorted in the field to ensure adequate numbers of juveniles, adequate numbers of adults of each sex (if sexually dimorphic) and to collect fish of similar size class. If the species chosen is not sexually dimorphic, an assumption is made that sex ratios in the field are equal and the fish should be randomly allocated to the tanks. It is possible to collect a subsample of these fish, dissect, and evaluate the sex distribution for additional information. Fish are transported back to the study site in containers with covers and under adequate aeration. If reference areas are not available, hatchery-reared fish can also be used if these fish are relevant to the study area.

At the study site, and preferably within the same day, fish are measured for length and weight, tagged if necessary and randomly assigned to each stream. The precision for these measurements is as described in the Technical Guidance Documents (Environment Canada 1995b; 1998; 2001) with increased precision when using small-bodied species. The optimal numbers of fish per stream are 20 adults of each sex and 30 juveniles. To ensure random allocation of fish to streams, 16 aerated buckets with reference water are set up. Fish are measured and placed into each bucket until each bucket contains one fish. This procedure is repeated until all buckets have five fish each. Buckets are then randomly allocated to the streams and the procedure is repeated until target sample sizes are attained in the streams. It is essential that this procedure be followed to ensure that the largest fish for example are not all allocated to the first few streams.

There are two possible fish allocation scenarios (Figure 5). Both adults and juveniles can be allocated to the same tanks if juvenile growth rates will be based solely on starting and ending length and weight measurements and adults are not cannibalistic (Figure 5a). If juveniles will be sampled during the course of the study for growth (perhaps juveniles are tagged and repeated measures are conducted) then allocating adults and juveniles to different tanks (or different areas within a tank if these areas can be physically separated)

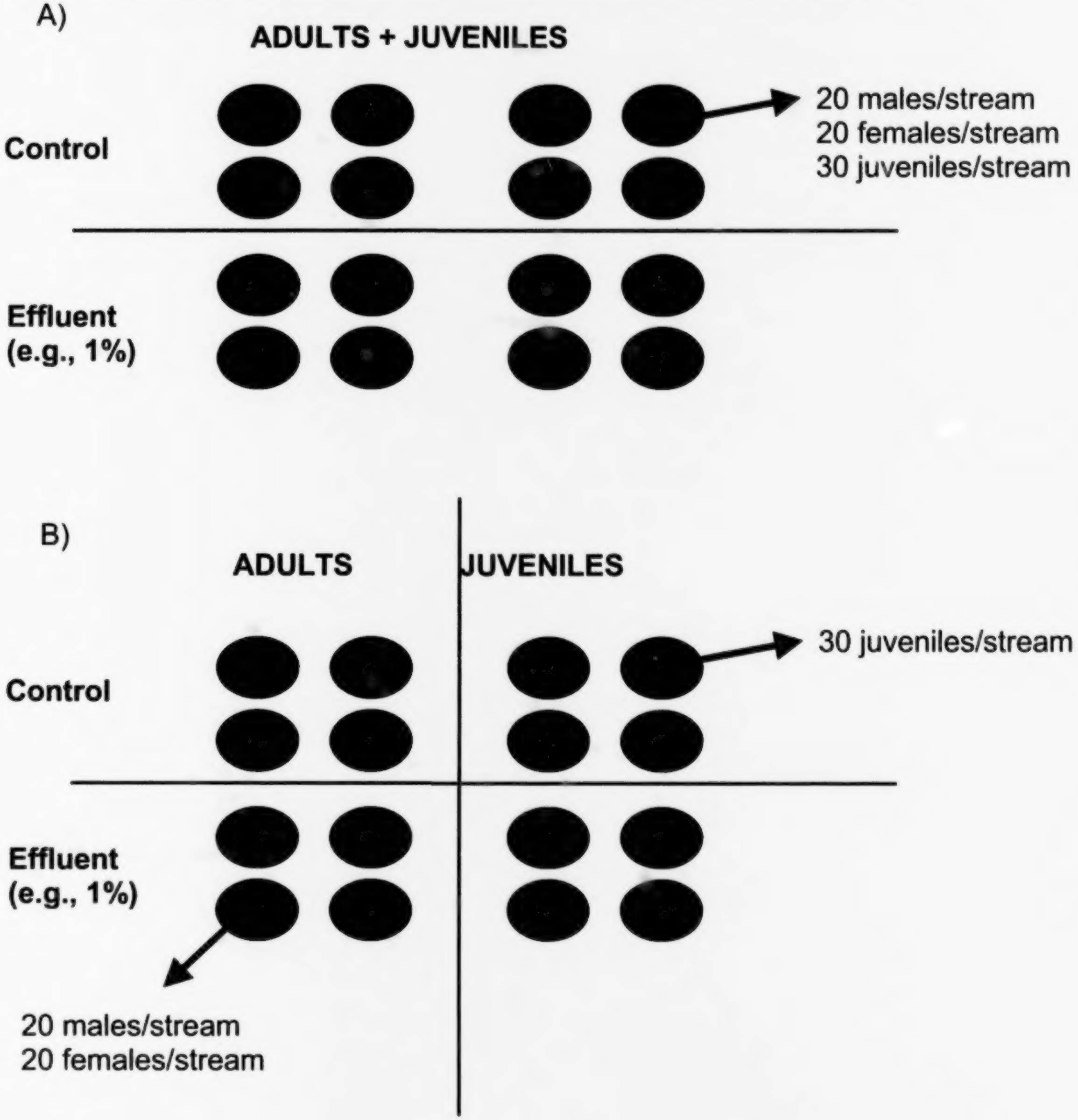

A)
ADULTS + JUVENILES
Control
20 males/stream
20 females/stream
30 juveniles/stream
Effluent
(e.g., 1%)
B)
ADULTS
JUVENILES
Control
30 juveniles/stream
Effluent
(e.g., 1%)
20 males/stream
20 females/stream

Figure 12-5 Fish allocation scenarios where A) adults and juveniles are allocated to the same stream or B) adults and juveniles are allocated to separate streams. For actual implementation, streams are randomly allocated to each treatment.

would be recommended to minimize adult capture stress (Figure 5b). Although Figure 5 indicates otherwise for ease of illustration, mesocosm streams are randomly assigned to each study treatment (control, effluent) and fish are randomly allocation to each stream. This ensures a randomized distribution to minimize stream differences due to factors such as stream position on the trailer.

After inoculation, fish are acclimated to artificial stream conditions for a minimum of 72 hours prior to effluent addition. Streams should be covered with netting to minimize fish loss due to escape or predation by birds. During the exposure period, which extends from 30 to 60 d, fish should be fed trout pellets at a rate of 4-6% total body weight per tank per day. The size of the pellets is dependent upon the size of the species used. In previous studies, trout pellets were crushed prior to feeding.

12.1.10.4.2 Maintenance and Monitoring

Daily maintenance for the duration of the study includes calibration of both water and effluent flows to each stream to ensure target effluent dilutions are attained; cleaning of screens; and feeding of fish. Screens are placed over the water outlet pipes on each stream to prevent fish loss (Figure 3c). These screens require daily cleaning to keep the overflow drains clear. Pump maintenance and changing of peristaltic lines due to clogging, algal growth, or leaking due to frictional wear over the pump roller heads, is also monitored and maintained on a daily basis.

Daily monitoring requirements for each stream include recording of any fish mortalities and monitoring of physical water chemistry including temperature, dissolved oxygen, conductivity, and pH. These measurements can be taken using YSI instruments or continuous recording equipment such as hydrolabs or thermisters. Dissolved oxygen is of critical importance especially in studies using effluents high in organic content (e.g., pulp mill effluents). Dissolved oxygen levels should be maintained above 60% (minimum) and preferable above 80% in all streams, using aeration if required.

12.1.10.4.3 Sampling and Analysis of Fish

The order of stream sampling should be random and preferably all streams should be sampled on the same day. If this is not possible, random sampling of the streams limits differences amongst streams within a treatment due to differences in the time of sampling over two days. All fish from each stream should be sampled before progressing to the next stream.

To sample a stream, water and effluent flows to that stream are stopped and water depth is reduced using screened siphons. Large rocks can be removed to improve ease of sampling and fish are collected using hand nets. It is easiest if males, females and juveniles are sorted into three separate buckets at the time of collection. Fish are held under cover and aeration until measurements are taken. An efficient 4-person measurement team should be able to process all fish from one stream in half an hour on the average.

The measurements required and level of precision are the same as those outlined in the Technical Guidance Documents (Environment Canada 1998; 2001). Sample preparation, laboratory analysis and QA/QC procedures for mesocosm studies are the same as those required for field fish surveys.

12.1.10.5 Data Assessment and Interpretation

12.1.10.5.1 General Requirements

For mesocosm studies, Data Assessment and Interpretation follows each monitoring or assessment phase in a similar way as for field surveys. In Data Assessment and Interpretation, the following questions are answered:

- is there an effect?
- is the effect mine or mill related?
- is the magnitude and extent of the effect known? and,
- is the mine or mill related cause of the effect known?

For the purposes of the EEM program, an effect on fish is defined as "a statistical difference between fish population measurements taken in an exposure area and reference area". An overview of data analysis and interpretation for mesocosm studies is presented here. However, Environment Canada (1998; 2001) should be consulted for further details.

Following Data Assessment and Interpretation, the mines will submit an interpretative report to the Authorization Officer (AO). This report will include information such as description of any changes from the submitted study design, the identification of any effects, the conclusions of the study, and a timeline outlining when the next monitoring phase will be conducted.

12.1.10.5.2 Statistical Analysis of Parameters

To determine whether or not there is an effect on fish exposed to effluents in mesocosm streams, statistical analysis of the data are conducted as suggested in Tables 3 and 4. The biological variables will be used to derive the following parameters:

Energy Use (Growth and Reproduction)

- growth: size of juveniles at end of the exposure period compared to starting measurements
- reproductive investment: gonad weight (as gonad weight relative to body weight; gonad weight relative to length), fecundity (as fecundity relative to body weight; fecundity relative to length)

Survival

- percentage surviving in the streams at the end of the exposure period

Energy Storage (Condition)

- condition (body weight relative to length)
- liver weight (as liver weight relative to body weight and liver weight relative to length).

Sex differences in growth rate, body weight, condition factor, gonadal size and liver size are common due to differences in overall energetic requirements between male and female fish. Therefore, for all parameters, sexes should be treated separately when estimating variability. Immature or juvenile fish should also be treated separately. The analyses that are suggested in Tables 3 and 4 are preferred. However, other analyses for specific parameters can been conducted depending upon the variability of the data set. For example, in mesocosm studies individuals of a similar size class are selected for placement in the streams resulting in reduced variability in endpoints compared to those measured in field surveys. Thus the range over which regressions such as ANCOVAs are conducted for mesocosm data is small. ANOVAs on original organ weights (liver size, gonad size), and preferably ratio metrics (LSI, GSI), or condition factors are justified in this instance.

The order of data analysis is largely a matter of personal preference, but the recommended methods are divided into five stages: a) preparing the analyses, b) initial summary statistics, c) ANOVA analyses, d) ANCOVA analyses, and e) power analyses. Details for these stages are available in Environment Canada (1998; 2001). Statistical considerations specific to mesocosm data sets are stated below.

Table 12-3 Effects on fish. Biological endpoints recommended for determining effects as designated by statistically significant differences between exposure and reference streams.

Type of Response	Effect Endpoint	Statistical Procedure
Survival	Juvenile and Adult Survival	ANOVA
Energy Use (Growth)	Juvenile Growth (change in length over time)	ANOVA (nested)
	Juvenile Growth (change whole animal wet weight over time)	ANOVA (nested)
(Reproduction)	Relative gonad size (gonad weight against body weight) - Adults	ANCOVA
Energy Storage (Condition)	Juvenile and Adult Condition (body weight against length)	ANCOVA
	Relative liver size (liver weight against body weight)- Adults	ANCOVA

Table 12-4 Biological endpoints to be used for supporting analyses. These analyses are for informational purposes and significant differences between exposure and reference streams are not necessarily used to designate an effect.

Type of Response	Supporting Response Variable	Statistical Procedure
Energy Use (Growth)	Adult Body weight (whole)	ANOVA (nested)
	Adult Length	ANOVA (nested)
(Reproduction)	Relative gonad size (gonad weight against length)- Adults	ANCOVA
	Relative fecundity (# of eggs/female against body weight)	ANCOVA
	Relative fecundity (# of eggs/female against length)	ANCOVA
Energy Storage (Condition)	Relative liver size (liver weight against length)	ANCOVA
	Relative egg size (mean egg weight against body weight)	ANCOVA

12.1.10.5.3 Nested ANOVA Analyses

For any variable measured once on a whole replicate stream, the statistical design is a simple t-test, comparing exposed and control treatments (control/impact design). The mesocosms or streams are replicates for the treatments. However, all biological and biochemical endpoints are measured on replicates at a lower level - either individual fish or composites of several fish within streams. A nested ANOVA is used to analyze these variables (Table 5). If the variance among replicate streams within a treatment (Error I) is large relative to the variance among fish or composite samples within streams (Error II), then the nested ANOVA is effectively a t-test comparing treatments, with the stream means as replicate observations. However, if Error I is not large relative to Error II (e.g. p>0.25), then the two error terms can be pooled to increase the power of the test comparing treatments (i.e. individual fish or composite samples rather than streams are used as replicates).

Table 12-4 ANOVA Table for Nested Design, with Two treatments, Four replicate Streams per treatment, and *n* Fish or Composite Samples Within Each Replicate Stream

Source		Df	F-test
Treatment	Tr	(Tr-1) = 1	MS (Tr) /MS (Error 1)
Streams within treatments	Error 1	Tr(r-1) = 6	MS (Error 1) /MS (Error 2)
Fish or composite samples within streams	Error 2	T(n-1) = 232	

Where: df = degrees of freedom; MS Mean Square
Tr = number of treatments, r = number of replicate tanks per treatments
T = number of tanks, n = number of fish per tank

12.1.10.5.4 ANCOVA Analyses

Most parameters are estimated as regressions of one variable on another using Analysis of Covariance (ANCOVA). ANCOVA combines the features of analysis of variance and regression, and can be used to compare regressions among multiple classifications (i.e. sampling areas) (Snedecor and Cochran 1980). The mean squared error (MSE; model variance, $s^2 = SD^2$ of the residuals from the regression]) of the ANCOVA model should be used to provide an estimate of variance of the fish parameters, when reasonably large samples (e.g., including 20 or more fish) are used. However, in mesocosm studies, the fish are chosen so that they are all within a comparatively narrow size/age range. Thus, it may not be necessary to use ANCOVA to factor out the influence of the size/age covariate on the endpoints. If limited fish availability necessitated the use of a wider size/age range, then ANCOVA comparisons might be required for some of the endpoints to factor our size/age effects.

12.1.10.6 Data QA/QC and Analysis

The importance of ensuring data quality can not be overstated. There are basics related to study design, consistency of methods and measurements, definition of protocols and procedures, which are outlined in detail in Environment Canada (1998; 2001).

Data should be tested for validity of statistical assumptions including normality and homogeneity of variance before applying parametric procedure as described in Environment Canada (1998; 2001). However, most univariate normal distribution-based statistical based methods are quite robust and can support moderate violations of the assumptions. Transformation of original data will help normalize the data or homogenize the variances. Logarithmic transformation are often preferred because most biological measures are considered to operate on a log or exponential scale (Peters, 1983). If the transformations are unable to produce data that meet the assumptions, then non-parametric statistics are used.

When designing experiments, it is also desirable to ensure that replicates are randomly allocated to different treatment levels such that the responses of each replicate is independent of other replicates. This element of randomness provides some assurance that observed differences in responses among treatments results from treatment-effects and not for some other reason. Randomly allocating replicates to different treatment levels is a relatively easy procedure when conducting manipulative experiments (e.g., mesocosm studies), but is less obvious for observational field studies. In natural environments, the high degree of spatial heterogeneity and the challenge of obtaining true replicates makes it difficult for biomonitoring studies to predict quantitatively the impacts of complex effluents on biota (Buikema and Voshell 1993, Forbes and Forbes 1994). Although biomonitoring provides important observations and insights into patterns of effluent effects on biota, inferential hypothesis testing through field assessments is often impossible (Stewart-Oaten et al. 1992; Cooper and Barmuta 1993). Experimentation using artificial streams is desirable in that they provide control over relevant environmental variables and true replication of treatments (McIntire 1993, Culp and Podemski 1996). Thus, causality (i.e., cause and effect) can be assigned definitively.

12.1.11 Benthic Invertebrate Community Monitoring Effects Assessment

12.1.11.1 Study Design

12.1.11.1.1 Species Selection

The benthic community that is established in the mesocosm streams is representative of that found in the reference field areas. Benthic samples are collected from the river and the entire community is inoculated into each mesocosm stream. Culp et al. (1996, 2001), Podemski (1999) and Culp et al. (2000a) have conducted studies that compared the community structure of benthic invertebrates in the mesocosm streams to that of field communities at reference areas. No ecologically significant differences in structure were observed, illustrating the effectiveness of the inoculation procedures and the suitability of mesocosms for testing effluent effects on environmentally relevant communities of benthic invertebrates.

The effectiveness of artificial streams in assessing effluent effects on environmentally relevant riverine communities was further corroborated in mesocosm studies on the Athabasca River (Culp and Podemski 1996; Culp et al. 2001). In these studies, effluent exposure significantly increased algal biomass and benthic insect abundances in the mesocosm streams. These results corresponded with in-river trends where periphyton biomass and benthic insect abundances increased at field sites downstream of the effluent (exposure areas) relative to upstream reference areas. These artificial stream studies substantiated speculation from earlier field studies which attributed higher abundances downstream of the outfall to pulp mill effluent-induced nutrient enrichment (Anderson 1989, Anderson 1991). Improved understanding of the mechanisms by which pulp mill effluent exposure affects the river ecosystem was achieved because the use of the stream-side mesocosms allowed the separation of confounded environmental factors at a community level (Culp and Podemski 1996).

12.1.11.2 Endpoints

12.1.11.2.1 Biological Endpoints

The objective of the benthic invertebrate component of the EEM programs is to determine if the mine or mill effluent has modified the benthic invertebrate community (i.e., is there an effect?). The mesocosm system can be used to assess the effects of effluents on a variety of benthic endpoints including those of community structure and process, as well as the growth and fecundity of selected community members (Culp et al.

1996, Culp and Podemski 1996, Podemski and Culp 1996, Culp et al. 2000, Culp et al. 2001). Community-level studies incorporate greater environmental realism (i.e., are more accurate in terms of incorporating the modifying effects of interspecies interactions on overall stressor effects), are better able to identify those components within the community most sensitive to a given effluent, and integrate a variety of ecological processes (e.g., competition and food web dynamics) that cannot be simulated in single species studies (Forbes and Forbes 1994). Furthermore, multispecies tests are a very efficient way to obtain information on the relative sensitivity of a large number of species in a single study. In mesocosm studies, we have been able to address community-level questions relating to the potential effects of varying concentrations of effluent on primary (attached algae) and secondary (benthic insects) producers. Specific endpoints measured in these studies included periphyton growth, the growth of selected insect taxa, algal and invertebrate community structure, and contaminant concentrations in algae (Table 1) (Culp and Cash 1995, Culp et al. 2000, Culp et al. 2001, Dubé et al., 2001a).

All endpoints used to assess effluent effects on benthic invertebrates in EEM field surveys can be measured in mesocosm studies. For EEM, an effect on the benthic invertebrate community is defined as a statistical difference between benthic invertebrate community measurements taken in an exposure area and a reference area, or a statistical difference between measurements taken at sampling areas in the exposure area that indicate gradually decreasing effluent concentrations, for any of the following key descriptors:

- Total invertebrate density
- Taxon (i.e., Family) Richness
- Simpson's Diversity Index
- Bray-Curtis Index (summarizes the overall difference in community structure between reference and exposed treatments in a single number and requires no preconceived assumptions about the nature of the community (Taylor and Bailey 1997)).

Evenness, taxon density, proportion or presence/absence are recommended to allow for the interpretation of effects (Environment Canada 2001). These descriptors are summary metrics selected to encompass the range of responses which may result from effluent, including changes in productivity, species composition and biodiversity. Many other benthic invertebrate descriptive metrics are available in the literature (for a review see Resh and McElravy 1993) and may be used, if applicable, on a site-specific basis to aid in the interpretation of effects determined with the key descriptors listed above.

12.1.11.3 Timing and duration of effluent exposure

Mesocosm studies using benthic invertebrates can be conducted at any time from early spring to late fall. The primary limiting factor is temperature with temperatures below -5 °C prohibitive to mesocosm operation due to freezing of water lines. Studies should be conducted during periods when field communities are under maximal effluent exposure for improved environmental realism. Studies should also be conducted at the time of year when the benthic invertebrate diversity is highest and water temperatures are conducive to growth. If historical data exists, it would be useful to examine the data and, if appropriate, conduct the study during similar periods so that the studies can be compared. Similarly, subsequent monitoring should be conducted during similar periods of the year to be comparable.

The duration of the mesocosm studies for benthic invertebrate programs is typically 30-45 days which includes a 7-12 day inoculation period for primary producers and a 25-30 day effluent exposure period. All community endpoints can, and have, been measured over this study duration. Effluents are collected daily or every second day during the studies and should be representative of normal industrial operating conditions.

12.1.11.4 Benthic Invertebrate Methods

12.1.11.4.1 General Set-up

General set-up requirements specified in Sections 12.1.3 and 12.1.6 also apply for benthic invertebrate studies. The recommended total flow of reference water or of dilution water and effluent into each stream is 2 L/min. This flow rate results in a 2 hour volume turnover time for each stream (stream volume 227 litres). Stream volume turnover rates of 4 hours have been used to reduce effluent requirements. However, 4 hour turnover times are the maximum. Longer turnover rates are not recommended due to effects on stream temperature and dissolved oxygen. For invertebrate applications, current velocity in each stream is created by a belt-driven propeller system as described in Section 12.1.3. Current in each stream is established based on site-specific conditions but is normally set between 10-20 cm/sec.

Prior to placement of substrate into the streams, five sampling bags (0.1 m^2, 500 μm mesh) are installed on the stream bottom. These bags are lifted at the end of the experiment resulting in five subsamples per stream. A standardized benthic environment is created in each stream to simulate the dominant, environmentally relevant habitat type. To date, mesocosm studies using benthic invertebrates have focused on riverine habitats where riffle substrate is dominant. The bottom of each stream and the sampling bags are covered with an 8 cm layer of washed gravel (1-2 cm diameter stones). The gravel is then left to colonize for a 7 to 12 day period to allow sufficient algal and microbial growth. Only water delivery to the mesocosm streams occurs (2 L/min) during this colonization period. The duration of the colonization period is site-specific and depends upon temperature and colonization of algae from the river into the system.

Following the algal colonization period, existing benthic communities are transplanted to the stream mesocosms. This protocol is determined on a site-specific basis. The following is an example of inoculation protocol for a riverine, riffle habitat with a large cobble/gravel dominated substratum:

1. Large cobbles (~x surface area = 535 cm^2) are randomly selected from the river at the reference area for placement into the mesocosm streams with their associated periphyton and invertebrate biota. These cobbles provide additional substrate and also stock the streams with a natural community of periphyton and benthic invertebrates.
2. During collection, the stones are enclosed with a 0.1 m^2 U-net (500 μm mesh) (Scrimgeour et al. 1993), carefully lifted from the streambed and placed into a container with river water. In addition, the gravel substratum beneath the stones is gently disturbed (to a depth of 5 cm) to collect any invertebrates under and around the base of the stone. The container is carefully transported to the mesocosm so that the periphyton and invertebrates associated with the stone are not dislodged. The cobbles are randomly placed in the artificial streams orientated to the water flow similar to their original orientation in the natural environment. This process continues until the appropriate density (e.g., 10 large cobbles/artificial stream) of large cobbles is reached based on natural substratum composition. Other samplers may be used to inoculate the streams with benthic communities as outlined in Environment Canada (1998; 2001). The objective however, is to establish ambient densities in the mesocosm streams. To determine the number of river samples to place in the streams, the mesocosm stream area (0.9 m^2) is divided by the area of the selected sampler. This random allocation of subsamples from pooled invertebrate community samples ensures that the initial invertebrate composition is similar amongst streams and limits the amount of variability in community composition among mesocosms attributable to the pattern of species introduction. (Wrona et al. 1982)
3. An additional series of benthic samples (biota only) are collected from the reference area and are pooled into a common container to estimate initial composition and density. Sub-samples (Marchant 1989) are then removed and randomly inoculated into each artificial stream until the density of invertebrates approximates ambient densities.

Standardization of inoculation protocols is important because the sequence through which species from a common source pool are added to mesocosms can produce large differences in community structure; these dissimilarities are unrelated to the intended study treatments (Drake et al. 1996). The test community includes multiple trophic levels constructed from random samples from the study river. Consequently, measurements of community-level endpoints, such as species composition, community production/respiration, and decomposition, are possible. These community-level measures integrate both the direct effects of stressors and, importantly, indirect ecological effects that cannot be simulated in single species or simple food chain systems (e.g., competition-mediated shifts in community structure, or biomagnification) (Carlisle 2000).

Following biotic inoculation, the mesocosm system is left to acclimatize for 24-48 hours (water flow only) prior to commencement of effluent delivery to the treatment streams.

12.1.11.4.2 Maintenance and Monitoring

Maintenance and monitoring is as described for the fish in section 12.1.10.4.2. Streams are inspected daily and drain screens are brushed clean to prevent clogging and overflow of the streams. Motors are lubricated with non-toxic, Permatex Superlube ® (food grade USDA H1) every 3-6 d to prevent seizing of the propeller shafts. The water delivery lines are flushed regularly to remove silt and sand deposits. Water and treatment solution flows to each stream are calibrated regularly. Effluent delivery tubing also requires daily inspection for blockages and wear from the pump heads, and changed as required.

Daily monitoring requirements for each stream includes monitoring of physical water chemistry including temperature, dissolved oxygen, conductivity, and pH. These measurements can be taken using YSI instruments or continuous recording equipment such as hydrolabs or thermisters. Dissolved oxygen is of critical importance especially in studies using effluents high in organic content (e.g., pulp mill effluents). Dissolved oxygen levels should be maintained 60% (80% preferred) in all streams using aeration if required.

12.1.11.5 Sampling and Analysis of Algae and Benthic Invertebrates

Sampling of algal communities in the mesocosm streams is considered a supporting measure to assessing effects of effluents on benthic invertebrate communities. However, the information for sampling and collection is presented here as these samples must be collected prior to sampling of benthic invertebrates to ensure sample integrity. Sampling of streams is done through randomized selection to minimize differences due to the time of sampling amongst replicate streams within the same treatment. Periphyton samples are collected once at the end of the exposure period, and a composite sample of three rocks per stream are sampled in order to reduce variability. Rocks are sampled by using a scalpel to remove periphyton from within a template of known area (e.g., 9.6 cm^2). Periphyton samples are placed in vials and held on ice until frozen later the same day. In the laboratory, each sample is homogenized, partitioned into two parts, and each portion filtered through a GF/C filter. Chlorophyll *a* concentration is determined by extracting the filter and retained material in an 80 °C bath of 90% ethanol for 5 min, then measuring fluorescence with a Turner Designs, model 10 series fluorometer. Ash-free dry mass (AFDM) is determined by weighing the sample after drying for 24 h at 105 °C, then combusting the filter at 500 °C and determining the weight loss upon ignition.

For benthic invertebrate sampling, streams are also sampled in a random selection to minimize differences due to the time of sampling amongst replicate streams within the same treatment. Invertebrate samples are collected by lifting the sampling bags within each stream which results in five subsamples per stream. Subsamples are then washed through a 500 μm mesh sieve. Field sieving is required immediately after sample retrieval and before preservation as described by Environment Canada (1998; 2001). The recommendation for sieve and/or mesh size for all freshwater mesocosm applications is

500 □m. In freshwater, macroinvertebrates are defined as those retained by mesh sizes of 200-500 µm (Slack et al. 1973; Weber 1973; Wiederholm 1980; Suess 1982), although immature life stages of some taxa may be smaller and some adult life stages may be larger. Note that these mesh sizes are applicable to all equipment used in the field and laboratory (i.e., both the Nitex mesh on the benthic samplers and sieving apparatus). In some site-specific circumstances it may be desirable for the samples to be screened for smaller organisms by using a smaller sieve size. Some examples of situations where the use of a smaller mesh size (less than 500 mm) may be appropriate include the following:

1) for comparative purposes if historical benthic surveys for the system under investigation utilized smaller mesh sizes, or
2) if sampling needs to be conducted, for logistic reasons, at times when organisms are very small Rees (1984), Barber and Kevern (1974) and Jonasson (1955) provide information on seasonal effects of mesh size.

In these aforementioned cases, it is highly recommended that a stack of screens be used which minimally have the mandatory sieve sizes and then any other smaller sizes which are appropriate. This procedure simultaneously allows site-specific concerns to be addressed and fulfil the EEM objectives by allowing for national or regional comparisons to be conducted on the standardized mesh sizes.

After sample sieving, invertebrates are preserved immediately in 10% formalin. In the laboratory, samples are sorted and identified. In previous mesocosm studies, the contents of three randomly chosen sampling bags were sorted from each stream. The recommended level of taxonomic identification is *family*. All summary statistics and descriptive metrics should be calculated and reported at the family level for submission for EEM interpretative reports. Organisms that cannot be identified to the desired level of taxonomic precision should be reported as a separate category in the fundamental data set. It is recommended that investigators use taxonomic keys appropriate to the geographic region of study; a detailed list of taxonomic references for various groups of freshwater organisms is provided in Environment Canada (1998; 2001)

12.1.11.6 Data Assessment and Interpretation

Use of mesocosms as a monitoring alternative for assessing effluent effects on the benthic invertebrate community largely follows the guidance of Environment Canada (1998; 2001) for data assessment and interpretation in field surveys as the endpoints measured to assess effects are the same. The only difference in data assessment is replication per treatment in mesocosm studies is n=8 for control/impact designs rather than n=5 as recommended for field surveys. Data Assessment and Interpretation follows each monitoring phase with the objective of answering the following questions:

- is there an effect?
- is the effect mine or mill related?
- is the magnitude and extent of the effect known? and,

- is the mine or mill related cause of the effect known?

During the effects assessment, a significant difference between reference and exposure areas in any of the following key descriptors is to be interpreted as an effect on the benthic invertebrate community.

1) Total invertebrate density
2) Taxon (i.e., Family) Richness
3) Simpson's Diversity Index
4) Bray-Curtis Index

Evenness, taxon density, proportion or presence/absence are recommended to aid in the interpretation of effects (Environment Canada 2001). Calculation of these metrics is described in detail in Environment Canada (1998; 2001). Details on recommended statistical analyses, data QA/QC, and reporting requirements is as outlined for field surveys with the qualification that a mesocosm treatment level is equivalent to an exposure area and replicate artificial stream is equivalent to a field replicate station.

12.1.12 Supporting Measurements

12.1.12.1 Water Chemistry Endpoints

In mesocosm studies, it is essential that stream differences in physical water chemistry be minimized as much possible to ensure effluent-related effects are not confounded. Streams are monitored daily for temperature, dissolved oxygen, pH, and input flow for water and effluent. Current velocity should be measured at the start and end of each study, especially for benthic invertebrate studies. The distribution of water velocities in the streams is characterized using one of many brands of a velocity meter. In previous studies, mean velocity in the mesocosms (above stones) ($x = 0.26 \pm 0.01$ m/sec, $n= 150$), was similar to water velocity measured above stones at a similar water depth in the field (water velocity $x = 0.26 \pm 0.01$ m/sec, water depth $x = 24.8 \pm 0.72$ cm, $n = 30$) while the study was being conducted (Podemski 1999).

Water temperature can be measured using continuous temperature loggers. In the latter instance, water temperature can be monitored with data loggers placed in a stream and the head tank. Temperatures in the head tank reflect the temperature of incoming river water. In previous studies, a comparison of data from these 2 thermograph locations indicated that the 2 h hydraulic residence time in the streams resulted in slight heating or cooling of water in the streams depending upon ambient air temperatures (Culp and Podemski 1996). For example, over a 3 d period, the streams were cooler at night and warmer during the day as compared to the incoming river water. The maximum instantaneous difference between water temperature in the river and the streams was $< 5\ ^{0}$C.

A repeated measured (RM) design is appropriate for analyzing physical and chemical water quality variables that are measured more than once (Table 6), since measurements are made on the same replicates at each time (Paine, 1996). The RM design is useful for evaluating the effectiveness of the mesocosm approach in controlling all variables except those related to exposure to the effluent. Normally, in an impact or effects assessment, the analyses of the physical and chemical variables are considered supplementary or secondary to analyses of the biological and chemical endpoints (i.e., effects measures) and are relegated to an appendix.

Table 12-6 ANOVA Table for Repeated Measure (RM) Design with Two Treatments, Four Replicate Streams per treatment, and t Measurements Times

Source		Df	F-test
Treatment	Tr	(Tr-1)	MS (Tr) /MS (Error 1)
Streams within treatments	Error 1	Tr(r-1)	MS (Error 1) /MS (Error 2)
Times	T	(t-1)	MS (T)/MS (Error 2)
Treatment x Times	Tr x T	(Tr-1)(t-1)	MS (Tr x T) /MS (Error 2)
Streams within treatments x Times	Error 2	Tr(r-1)(t-1)	

Where: df = degrees of freedom; MS Mean Square
Tr = number of treatments, r = number of replicate within treatments
t = number of repeated measurements

12.1.13 Exposure Endpoints

In addition to assessment of daily changes in physical water chemistry amongst all of the mesocosm streams, water samples should also be collected on a less frequent basis for analysis of general chemical parameters. It is recommended that samples be collected from the reference water head tank, from the full strength effluent, and from each mesocosm stream once at the end of the study. Samples should be analyzed for parameters as outlined in Environment Canada (water quality monitoring parameters for metal mining, supporting parameters for pulp and paper) (1998; 2001). Often this information proves invaluable to confirm effluent dilutions in the treatment streams. For example, in freshwater systems exposed to pulping effluents, sodium is a relatively conservative ion that appears in high concentrations. By comparing Na^+ concentrations in the reference and exposure streams, the concentration of effluent in the streams can be verified.

Rhodamine dye has also been used on several occasions in mesocosm studies, following sampling of biota, to confirm effluent concentrations in the treatment streams. A concentrated solution of rhodamine is established in a storage tank and delivered to the treatment streams through the existing effluent lines. After a 24 hour period of turnover, water chemistry samples are collected in the rhodamine head tank and in each effluent treatment stream to calculate exposure concentrations.

The primary advantage of mesocosm systems is that they offer controlled exposure conditions. Thus chemical measurements to confirm exposure in biota are not normally required. However, these measures can provide valuable supporting information for site-specific assessments. For example, Dubé et al., (2001b) measured resin acids (i.e., dehydroabietic acid) in the bile of mummichog exposed to environmentally relevant concentrations of a mill effluent (3%) and compared these measures to control fish. Results showed significantly higher concentrations of DHAA in fish which confirmed that DHAA was a suitable tracer to confirm exposure to secondary treated pulp mill effluent where the mill furnish has a high proportion of softwood. These results also showed effluent-related differences between control and exposed fish.

Metal levels in indicator tissues have also been used as supporting information in mesocosm studies examining mine effluent effects on fish. Composite muscle tissue from juvenile Atlantic salmon showed significantly higher metal levels in fish exposed to mining effluent for 39 days compared to control fish (Dubé et al., unpublished). The metals measured in tissues were consistent with those detected in the effluent stream. The indicators and the tissues to measure metal exposure will vary with mine type and species being used. In general, gills, liver and kidney have the greatest potential for estimating exposure and bioavailability of metals. For small species (< 10 cm), whole body levels can be examined, or levels in the carcass after removing the digestive tract. Mercury is the only element that has been found to accumulate in muscle tissue, so if mercury is a contaminant of concern, dorsal muscle tissue should be analyzed. Blood and bone tissue may reflect exposure to lead, and might be considered if lead is the primary element of concern (Hodson et al., 1984). Bone concentrations are expected to be most indicative of long term metal exposure, while blood concentrations are indicative of short term exposure (AETE 1998).

12.1.14 Other Considerations

The guiding principles of the EEM program are that the program be scientifically defensible, cost-effective and provide flexibility for site-specific requirements. In order to keep the EEM programs focused and cost-effective, screening criteria were developed in the pulp and paper EEM program to ensure that only those monitoring techniques that were environmentally relevant, interpretable, and scientifically defensible were recommended in the EEM program (Table 7). Artificial stream systems satisfied all of the stated criteria with the following exceptions. For detecting the effects of effluents on fish, artificial stream studies do not examine multi-species interactions at a community of population level. However, it should also be recognized that the existing field surveys for fish also do not examine multi-species interactions at the community or population level.

Other limitations to application of artificial stream systems can include the lack of appropriate access and infrastructure requirements at some monitoring sites, and the restriction to medium term (i.e., weeks to months) monitoring such that long-term bioaccumulation processes cannot be examined.

Table 12-7 Evaluation criteria for EEM monitoring alternatives (after Dubé et al., 2001a).

Evaluation Criteria	Description
1. Environmentally Relevant	• Effects on relevant organisms are measured in the receiving environment.
2. Interpretable	• Measured endpoints are relevant to the detection/quantification of effects by pulp and paper (or mining) effluents on fish, fish habitat or the use of fisheries resources. • Amenable to consistent scientific evaluation and interpretation. • An approach for determining "an effect" (statistically significant difference between reference and exposed) has been demonstrated.
3. Scientifically Defensible and Manageable	• Acceptable test protocol exists, including QA/QC procedures. • Commercially available. • A validation study (e.g., pilot studies) has been completed.
4. Cost-effective	• The new technique provides equal or better quality data in comparison with currently accepted techniques. • Cost of the technique is similar to currently accepted techniques.

Cost-effectiveness of mesocosm application is critical for EEM application. Mesocosms can be an effective tool in a regulatory setting providing the system employed is cost-effective, and the design suited to the objectives of the program and the hypotheses to be tested. The costs for many components of a mesocosm study are the same as those for a field survey because the endpoints measured are the same. Laboratory analysis costs for example are the same. The main difference in cost between field surveys and mesocosm studies is in labor and use of the mesocosm facility itself. Labor costs and system transport costs depend on where the study is being conducted (i.e. site-specific) and the length of the exposure period. Transportation of the facility to the site, hook up of power and water, and colonization with biota are typically conducted within 4-5 days with a 4-person field crew and a mine electrician. Costs for system use and maintenance are supplier-specific. If suppliers have their personnel on-site for the duration of a 60 d exposure period then costs would be higher than a field survey because of supplier rates. Our approach has been to have crews for set-up and take-down, with day to day maintenance covered by mine personnel, or summer students, as these tasks take 1/2 day at most and do not require high levels of scientific expertise. As an estimate, mesocosm studies should fall within the range of field surveys + approximately 25%.

It must be recognized that mesocosms are alternatives to field surveys. If field surveys cannot be conducted then an alternative should be used to provide the required information. All efforts should be made to use alternatives in a cost-effective manner, as this is the most appropriate mechanism to attain quality data that is interpretable for confounded sites. If mesocosms are considered as an alternative for a particular mine, a further consideration to ensure cost effectiveness would be to incorporate more than one phase or more than one effluent (e.g., if joint studies between mines are possible) into the study design. For example, if several environmentally relevant levels of effluent exposure are included in the design, the objectives for Initial Monitoring and Focused Monitoring could be met within the same study.

12.1.15 References

AETE (Aquatic Effects Technology Evaluation Program). 1998. Summary and cost-effectiveness of aquatic effects monitoring technologies applied in the 1997 AETE field evaluation program. Canada Centre for Mineral and Energy Technology, Ottawa, ON, Canada.

Anderson, A.M. 1989. An assessment of the effects of the combined pulp mill and municipal effluents at Hinton on the water quality and zoobenthos of the Athabasca River. Report, Environmental Quality Monitoring Branch, Environmental Assessment Division, Alberta Environment, Edmonton, AB, Canada.

Anderson, A.M. 1991. An overview of long-term zoobenthic monitoring in Alberta rivers (1983-1987). Report, Environmental Assessment Division, Alberta Environment, Edmonton, AB, Canada.

Barber, W.E., and N.R. Kevern. 1974. Seasonal variation of sieving efficiency in a lotic habitat. Freshwat. Biol. 4:293-300.

Buikema, A., and J.R. Voshell. 1993. Toxicity studies using freshwater benthic macroinvertebrates. In: D.M. Rosenburg and V.H. Resh (eds.), Biomonitoring and benthic invertebrates. Chapman and Hall, NY, USA, pp.340-395.

Burt, A., Kramer, D.L., Nakasuru, K. and C. Spry. 1988. The tempo of reproduction in *Hyphessobrycon pulchripinnis* (Characidae), with a discussion on the biology of "multiple spawning". Environ. Biol. Fish. 22: 15-27.

Carlisle, D.M. 2000. Bioenergetic food webs as a means of linking toxicological effects across scales of ecological organization. J. Aquat. Ecosystem Stress and Recovery 7:155-165.

Coad, B.W. 1995. Encyclopedia of Canadian Fishes. Canada Museum of Nature and Canadian Sportfishing Productions Inc. 928p.

Cooper, S. D., and L. Barmuta. 1993. Field experiments in biomonitoring. In: Rosenberg, D.M., and V.H. Resh (eds.), Biomonitoring and benthic invertebrates. Chapman and Hall, New York, NY, USA, pp.399-441.

Culp, J.M., and K.J. Cash. 1995. Potential impacts of effluent on the Thompson and upper Fraser Rivers. DOE FRAP 1995-10, Fraser River Action Plan, Environmental Conservation Branch, Environment Canada, Vancouver, BC, Canada.

Culp, J.M., and C.L. Podemski. 1996. Design and application of a novel stream microcosm system for assessing effluent impacts to large rivers. In: Servos, M.R., Munkittrick, K.R., Carey, J.H., and G.J. Van Der Kraak (eds.), Environmental fate and effects of pulp and paper mill effluents. St. Lucie Press, Delray Beach, FL, USA, pp.549-555.

Culp, J.M., Podemski, C.L., Cash, K.J., and R.B. Lowell. 1996. Utility of field-based artificial streams for assessing effluent effects on riverine ecosystems. J. Aquat. Ecosystem Health 5:117-124.

Culp, J.M., and R.B. Lowell. 1998. Pulp mill effluent impacts on benthic communities and selected fish species in the Fraser River Basin. In: Gray, C., and T. Tuominen (eds.), Health of the Fraser River Aquatic Ecosystem, Vol. II. DOE FRAP 1998-11, Environment Canada, Ottawa, ON, Canada, pp.13-34.

Culp, J.M., Lowell, R.B., and K.J. Cash. 2000a. Integrating mesocosm experiments with field and laboratory studies to generate weight-of-evidence risk assessments for large rivers. Environ. Toxicol. Chem. 19:1167-1173.

Culp, J.M., Podemski, C.L., Cash, K.J., and R.B. Lowell. 2000b. A research strategy for using stream microcosms in ecotoxicology: integrating experiments at different levels of biological organization with field data. J. Aquat. Ecosystem Stress and Recovery 7:167-176.

Culp, J.M., Podemski, C.L., and K.J. Cash. 2001. Interactive effects of nutrients and contaminates from pulp mill effluents on riverine benthos. J. Aquat. Ecosystem Stress and Recovery, in press.

Drake, J.A., Huxel, G.R., and C.L. Hewitt. 1996. Microcosms as models for generating and testing community theory. Ecology 77:670-677.

Dubé, M.G. and J.M. Culp. 1996. Growth responses of periphyton and chironomids exposed to biologically treated bleached kraft pulp mill effluent. Environ. Toxicol. Chem. 15:2019-2027.

Dubé, M.G. 2000. Sources of contaminants in the bleached Kraft pulping process and their effect on the mummichog (*Fundulus heteroclitus*). Ph.D. Thesis, University of New Brunswick, Saint John, NB, Canada, 211p.

Dubé, M.G. and D.L. MacLatchy. 2000. Endocrine responses of an estuarine fish *Fundulus heteroclitus* to final effluent from a bleached kraft pulp mill before and after installation of reverse osmosis treatment of a waste stream. Environ. Toxicol. Chem 19:2788-2796.

Dubé, M.G. and D.L. MacLatchy. 2001. Identification and treatment of a waste stream at a bleached kraft pulp mill that depresses a sex steroid in the mummichog (*Fundulus heteroclitus*). Environ. Toxicol. Chem 20:985-995.

Dubé, M.G., Culp, J.M., Cash, K.J., Glozier, N.E., MacLatchy, D.L., Podemski, C.L., and R.B. Lowell. 2001a. Artificial streams for environmental effects monitoring (EEM): development and application in Canada over the past decade. Wat Quality Res J Can, in press.

Dubé, M.G., Culp J.C., MacLatchy, D.L., Gillis, G., Parker, R., Courtenay, S., and C. Gilman. 2001b. Utility of field-based artificial streams for assessing effects of pulp mill effluents on fish. J. Aquat. Ecosystem Stress Recovery, in press.

Environment Canada. 1992. Pulp and Paper Effluent Regulations, Canada Gazette, Part II, Volume 126, Number 11:1967-1997.

Environment Canada. 1995a. Further guidance for the invertebrate community survey for aquatic environmental effects monitoring related to Federal Fisheries Act Requirements. EEM 2, Ottawa, ON, Canada.

Environment Canada. 1995b. Further guidance for the adult fish survey for aquatic environmental effects monitoring related to Federal Fisheries Act Requirements. EEM 1, Ottawa, ON, Canada.

Environment Canada. 1997. Aquatic Environmental Effects Monitoring Requirements (revised EPS 1/RM/18). EEM/1997/1, Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch Ottawa, ON, Canada.

Environment Canada. 1998. Pulp and Paper Technical Guidance for Aquatic Environmental Effects Monitoring. EEM/1998/1. Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch, Ottawa, ON, Canada.

Environment Canada. 2001. Metal Mining Technical Guidance Document for Aquatic Environmental Effects Monitoring. EEM/2001/in press. Environment Canada, National EEM Office, Science Policy and Environmental Quality Branch, Ottawa, ON, Canada.

Forbes, V.E., and T.L. Forbes. 1994. Ecotoxicology in theory and practice. Chapman and Hall, New York, USA.

Fritz, E.S., Meredith, W.H., and V.A. Lotrich. 1975. Fall and winter movements and activity level of the mummichog *Fundulus heteroclitus*, in a tidal creek. Chesapeake Sci. 16: 211-215.

Gibbons, W.N., Munkittrick, K.R., and W.D. Taylor. 1998a Suitability of small fish species for monitoring the effects of pulp mill effluent on fish. 1. Response of spoonhead sculpin exposed to bleached kraft mill effluent. Environ. Toxicol. Chem., in press.

Gibbons, W.N., Munkittrick, K.R., and W.D. Taylor. 1998b. Suitability of small fish species for monitoring the effects of pulp mill effluent on fish. 2. Comparison of trout perch responses to white sucker exposed to bleached kraft mill effluent. Environ. Toxicol. Chem., in press.

Green, R.H. 1993. Application of repeated measures designs in environmental impact and monitoring studies. Austral. J. Ecology 18:81-98.

Heins, D.C. and F.G. Rabito Jr. 1986. Spawning performance in North American minnows: direct evidence of the occurrence of multiple clutches in the genus *Notropis*. J. Fish. Biol. 28: 343-357.

Hodson, P.V., B.R. Blunt, and D.M. Whittle. 1984. Monitoring lead exposure of fish. Pg. 87-98. *In:* Contaminant Effects of Fisheries. Cairns, V.W., P.V. Hodson, and J.O. Nriagu. John Wiley and Sons. New York, New York.

Jenkins, R.E. and N.M. Burkhead 1993. Freshwater fishes of Virginia. American Fisheries Society, Bethesda, ML, USA, 1079p.

Jonasson, P.M. 1955. The efficiency of sieving techniques for sampling freshwater bottom fauna. Oikos 6:183-207.

Kneib R.T., and A.E. Stiven. 1978. Growth, reproduction, and feeding of *Fundulus heteroclitus* (L.) on a North Carolina salt marsh. J. Exp. Mar. Biol. Ecol. 31:121-140.

Leblanc, J. et C. M. Couillard. 1995. Description de la période de reproduction d'un poisson sentinelle: le choquemort (Fundulus heteroclitus) de l'estuaire de la Miramichi Rapp. tech. can. sci. halieut. aquat. 2057: vii + 32p.

Leblanc, J., Coulliard, C.M., and J.F. Brethes. 1998. Modifications of the reproductive period in mummichog (*Fundulus heteroclitus*) living downstream from a bleached Kraft pulp mill in the Miramichi Estuary, New Brunswick, Canada. C.J.F.A.S., in press.

Lowell R.B., Culp J.M., and F.J. Wrona. 1995. Stimulation of increased short-term growth and development of mayflies by pulp mill effluent. Environ. Toxicol. Chem. 14:1529-1541.

Lowell R.B., Culp J.M., Wrona F.J., and M.L. Bothwell. 1996. Effects of pulp mill effluent on benthic freshwater invertebrates: food availability and stimulation of increased growth and development. In: Servos, M.R., Munkittrick, K.R., Carey, J.H., and G.J. Van Der Kraak (eds.), Environmental fate and effects of pulp and paper mill effluents. St. Lucie Press, Delay Beach, FL, USA, pp.525-532.

Lowell R.B., Culp J.M., and M.G. Dubé. 2000. A weight-of-evidence approach for northern river risk assessment: integrating the effects of multiple stressors. Environ. Toxicol. Chem. 19:1182-1190.

McIntire, C.D. 1993. Historical and other perspectives of laboratory stream research. In: G.A. Lamberti and A.D. Steinman (eds.), Research in artificial streams: applications, uses and abuses. J. North Amer. Benth. Soc. 12:313-323.

Nelson, J.S. and M.J. Paetz. 1992. The fishes of Alberta. The University of Calgary Press, Calgary, AB, Canada.

Paine, M.D. 1990. Life history tactics of darters (*Percidae: Etheostomatiini*) and their relationship with body size, reproductive behaviour, latitude and rarity. J. Fish Biol. 37: 473-488.

Paine, M.D. 1996. Repeated measures designs. Environ. Toxicol. Chem. 15:1439-1441.

Peters, R.H. 1983. The ecological implication of body size. Cambridge University Press, New York, USA, 329p.

Podemski C.L., and J.M. Culp. 1996. Nutrient and contaminant effects of bleached kraft mill effluent on benthic algae and insects of the Athabasca River. In: Servos, M.R., Munkittrick, K.R., Carey, J.H., and G.J. Van Der Kraak (eds.), Environmental fate and effects of pulp and paper mill effluents. St. Lucie Press, Delay Beach, FL, USA, pp.571-580.

Podemski C.L. 1999. Ecological effects of a bleached kraft pulp mill effluent on benthic biota of the Athabasca River. Ph.D. Thesis, University of Saskatchewan, Saskatoon, SK, Canada, 292 p.

Rees, H.L. 1984. A note on mesh selection and sampling efficiency in benthic studies. Mar. Poll. Bull. 15:225-229.

Resh V.H., and E.P. McElravy. 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates. In: Rosenberg, D.M. and V.H. Resh (eds.), Freshwater biomonitoring and benthic invertebrates. Chapman & Hall, NY, USA, pp.159-194.

Roberts, W.E. 1988. The sculpins of Alberta. Alberta Naturalist 18:121-127.

Scott, W.B. 1967. Freshwater fishes of eastern Canada. University of Toronto Press, Toronto, ON, Canada. 137p.

Scott, W.B. and E.J. Crossman, 1973. Freshwater fishes of Canada. Fish. Res. Board. Can. Bull. 184. Ottawa, ON, Canada. 966p.

Scrimgeour, G.J., Culp, J.M., and N.E. Glozier. 1993. An improved technique for sampling lotic invertebrates. Hydrobiologia 254:65-71

Shuter, B.J. 1990. Population-level indicators of stress. Amer. Fish. Soc. Symposium 8:145-166.

Slack, K.V., Averett, R.C., Geeson, P.E., and R.G. Lipscomb. 1973. Methods for collection and analysis of aquatic biological and microbiological samples. In: Techniques of water-resources investigations of the United States Geological Survey, Chapter 4A, Book 5, U.S. Department of the Interior, Geological Survey, Washington, D.C., USA, pp. 1-165.

Snedecor, G.W. and W.G. Cochran. 1980. Statistical methods, 7th ed. Iowa State University Press, Ames, Iowa, USA, 507p.

Stewart-Oaten, A., Bence, J.R., and C.W. Osenburg. 1992. Assessing effects of unreplicated perturbations: no simple solutions. Ecology 67:929-940.

Suess, M.J. 1982. Examination of water for pollution control. A reference handbook. Vol. 3. Biological, bacteriological and virological examination. Permagon Press, Oxford, England.

Taylor B.R., and R.C. Bailey. 1997. Technical evaluation on methods for benthic invertebrate data analysis and interpretation, final report. AETE project 2.1.3. Natural Resources Canada, CANMET, Ottawa ON, Canada, 26p.

Weber, C.I. 1973. Biological field and laboratory methods for measuring the quality of surface waters and effluents. EPA-670/4-73-001. U.S. Environmental Protection Agency, Cincinnati, OH, USA.

Wiederholm, T. 1980. Use of benthos in lake monitoring. J. Wat. Poll. Cont. Fed. 52:537-547.

Wrona, F.J., J.M. Culp and R.W. Davies. 1982. Macroinvertebrate subsampling: A simplified apparatus and approach. Can. J. Fish. Aquat. Sci. 39:1051-1054.

12.2 Use of Caged Bivalves as Alternative Monitoring Method

12.2.1 Introduction

A description of methods for caged bivalve studies is provided in this section, which includes detailed guidance on:

- background and general approach of caged bivalve studies;
- species selection;
- study design;
- variables to be measured;
- methods for implementing the study;
- data analyses;
- tissue concentrations; and
- reporting requirements.

12.2.2 Background

In October 2000, the National Environmental Effects Monitoring (EEM) Science Committee recommended the use of caged bivalve studies as a scientifically defensible alternative approach to a wild fish survey if it is not practical or technically feasible to conduct the fish survey (FS), according to the decision tree in Annex 1 of the Pulp and Paper EEM Technical Guidance Document (Environment Canada 1998: Figure 5.1). The Metal Mining Fish Subgroup also recommended the use of caged bivalves as an alternative method during the Metal Mining EEM multi-stakeholder consultation process.

During the first cycle (1993 to 1996) of the pulp and paper EEM program, the FS performed reasonably well in the freshwater environment, but a number of problems were encountered for marine and estuarine environments. In February 1997, a workshop was convened to discuss alternatives to the FS component of the EEM program (Courtenay et al. 1998). One of the suggested alternatives was caged bivalve studies. Subsequently, two pilot studies using caged bivalves were conducted for marine environments: in 1997 at Port Alice on Vancouver Island (Applied Biomonitoring 2000) and in 1998 at Pictou Harbour in Nova Scotia (Andrews and Parker 1999). Based on these two pilot studies and several other related studies using caged bivalves, the Science Committee and Metal Mining Fish Subgroup determined that an EEM study using caged bivalves can produce good quality data on effluent effects on fish that fit into the regulatory context.

This technical guidance describes procedures for conducting controlled experiments using caged bivalves suspended in the water column, to test for effects associated with industrial discharges, comparing measurements from an exposure area to reference areas. This approach is an alternative for the FS and may be considered for mines where the FS has been unsuccessful or impractical in past cycles of EEM, or where there may be study design issues, such as confounding influences, or safety concerns.

In November 2000, American Society for Testing and Materials (ASTM) approved a method for conducting environmental studies using caged bivalves (Salazar and Salazar 2000). The ASTM method serves as the basis for this technical guidance on conducting caged bivalve studies. A number of other studies using caged and wild bivalves were also considered in developing this specific guidance for applying this approach within the framework of the EEM program. The format of this guidance follows that used for the FS.

Bivalves are useful biomonitors for several reasons:

- bivalves are relatively non-mobile, such that exposure to contaminants is assured and representative of the exposure area;
- bivalves are abundant in many marine, estuarine and freshwater environments, and are relatively easy to handle and sample year-round;
- the biology of many shellfish species is well known and considerable research has been conducted on effects to shellfish from exposure to various anthropogenic and natural environmental stressors;
- several bivalve species have been shown to readily accumulate many chemicals from a variety of pathways (water, sediment, food) and show sublethal effects associated with exposure;
- bivalve growth is relatively easy to measure and has been shown to be as sensitive or more sensitive than mortality in other standard test species such as *Daphnia*, fathead minnow and rainbow trout (see Salazar and Salazar 2000); and
- bivalves are an important fisheries resource and both the Atlantic and Pacific regions support commercially valuable shellfish aquaculture industries, as well as commercial and recreational shellfish harvests.

Bivalves, such as oysters and mussels, have been used in Mussel Watch programs since the mid-1970s to monitor trends in chemical contamination and assess the effects of human activities on coastal and estuarine areas. Mussel Watch programs began in the U.S. (Goldberg et al. 1978) and have since become international in scope (Jernelov 1996).

12.2.3 Approach

The caged bivalve approach provides a number of advantages to the investigator in conducting a monitoring program. These include experimental control and realism, use of organisms naturally found in the study environment, and known exposure period. By using caged bivalves, rather than resident populations, the variability in biological measurements can be reduced by using individuals of similar size and environmental history, thereby increasing the discriminating power of the test (Salazar and Salazar 1995). A considerable number of caged bivalve studies have been conducted in Canada and the U.S., as well as other countries (see Salazar, in Stewart and Malley 1997).

The target endpoints for caged bivalve studies in EEM are primarily survival, growth and condition. Reproduction is considered a key endpoint for fish population monitoring, but currently there are no standardized methods to monitor this in caged bivalves. Once these methods are established, this document will be updated to include guidance on this. A tissue analysis (dioxins and furans at pulp and paper mills; mercury at metal mines) may be required, and caged bivalves can be used to meet this requirement. Other chemicals or metals may be used to assess bioaccumulation to aid in interpreting results or used during investigation of cause.

There are some limitations to using caged bivalve studies as surrogates for fish surveys. These include (from Salazar, in Stewart and Malley 1997):

- possible effects of transplanting causing stress to bivalves;
- loss of cages from acts of nature, inadvertent capture by moving vessels, vandalism; and
- costs associated with collection, sorting and deployment.

That being stated, the caged bivalve approach is considered to be a potential alternative in some cases where a wild fish survey cannot be successfully conducted.

12.2.4 Species Selection

Many bivalve species have been used for assessing chemical bioavailability or effects in marine, estuarine and freshwater environments.

Species selection for caged bivalve studies should be made carefully and consider at least the following questions:

- Are conditions at the exposure and reference areas similar to the natural habitat of the species in terms of tolerance limits for natural factors such as temperature, salinity, dissolved oxygen, and pH?
- Is the species naturally present in the area under evaluation?
- Is there documentation that the species can accumulate and/or be sensitive to the contaminants of concern?
- Is the life history of the species well known in terms of spawning cycle and life-stage requirements?
- Does the species have threatened or endangered status?

- Is an abundant supply of the species readily available?
- Is the species easy to handle in the field?

Ideally, species should be used that have wide geographic distributions so that test results can be compared across studies.

12.2.4.1 Most Commonly Used Taxa

The species most commonly used in field bioassays in Canadian waters and considered relevant for use in the EEM programs are described below and listed in Table 8. The temperature and salinity tolerance limits for each species are provided, along with information on age at maturity, spawning periods and general distribution within Canada. Other species, such as marine clams *(Mya arenaria, Macoma balthica*) and scallops, may also be suitable for some applications. However, the environmental requirements and sensitivity of alternative bivalve test species should be established before they are used in an EEM program.

Table 12-8 Suggested Taxa for Use in Caged Bivalve Studies for EEMs

Species and Reference	Temperature Range (°C)	Salinity Range (parts per thousand)	Reproductive Information	General Distribution in Canada
Marine and Estuarine Bivalves				
Mytilus edulis (blue mussel) (Freeman et al. 1994; Grout and Levings 2000; Mucklow 1996; Stewart 1994; Salazar and Salazar 2000; Newell 1989)	-1.5 to 25	5 to 33	Matures at 2 years of age. Spawning varies among populations; some low-level throughout year; first in early summer, second in the fall	Atlantic coast
Mytilus trossulus (Pacific blue mussel) (Freeman et al. 1994; Salazar and Salazar 2000; Skidmore and Chew 1985)	0 to 29	4 to 33	Matures at 2 years of age. Spawning from July to September	Atlantic and Pacific coasts
Crassostrea gigas (Pacific oyster) (Virginia Tech 2000; Waldock et al.1996)	4 to 24	25 to 35	Spawning July to August.	Pacific coast
Freshwater Bivalves				
Elliptio complanata (Eastern elliptio) (Beckvar et al. 2000; Day et al. 1990; Hinch and Green 1989; McMahon 1991; Metcalfe-Smith et al. 1996)	0 to 30	0 to 3	Age at maturity 6-12 years. Spawning occurs mostly June to July; some May to September.	Eastern Canada
Pyganodon (Anodonta) grandis (common floater clam) (Clarke 1973; Couillard et al. 1995a; 1995b; Malley et al. 1996)	0 to 30	0 to 3	Spawns mostly April to May, some to August	Interior and Eastern Canada
Anodonta kennerlyi (western floater clam) (BC Research Inc. 1998; Clarke, 1981; Stewart and Malley 1997; Williams et al. 1993)	0 to 30	freshwater	Spawning begins in early August.	Alberta and British Columbia
Sphaeriid clams (e.g., *Musculium securis, Sphaerium rhomboideum, S. striatinum*) (Hornbach et al. 1982; Mackie 1978a; 1978b; Mackie and Flippance 1983; Mackie et al. 1974; Stephenson and Mackie 1981)	10 to 25* (* optimal growth range)	freshwater	Life cycle generally #1 year; life histories of many species are well documented; reproductive effort can be quantified.	Widely distributed in Canada

(Format modified from Salazar and Salazar 2000.)

12.2.4.2 Marine and Estuarine Bivalves

Mytilus sp. is an intertidal bivalve that has been successfully used in caged studies for marine and estuarine environments since the late 1970s. *Mytilus* has been used extensively in the ongoing International Mussel Watch Project to monitor trends of chemical contamination and assess the effects of human activities on coastal and estuarine areas in North America and around the world (O'Connor 1992; Jernelov 1996). The blue mussel, *M. edulis*, and *M. trossulus* (Pacific blue mussel) are found on the Atlantic Canada coast, whereas only *M. trossulus* is found on the Pacific Canada coast. These two species may be easily confused where they co-occur on the Atlantic coast (Freeman et al. 1994; Mucklow 1996). *M. edulis* was successfully used for the caged bivalve pilot studies for pulp and paper EEM on the Pacific coast (Applied Biomonitoring 2000) and the Atlantic coast (Andrews and Parker 1999) and has otherwise been widely used as a biomonitor (Stewart and Malley 1997; Salazar and Salazar 2000). *M. trossulus* has been successfully used in a number of caged bivalve studies in Pacific coastal waters (Stewart and Malley 1997).

Crassostea gigas is an oyster species that has been used in transplant studies in marine and estuarine studies (Waldock et al. 1996). The shells of *C. gigas* are usually more difficult to measure because of their irregular shape and protrusions.

12.2.4.3 Freshwater Bivalves

Unionid bivalves (such as *Elliptio complanata* and *Pyganodon grandis* (formerly *Anodonta grandis*)) are freshwater unionid mussels that have been used in a number of caged studies to examine water column and sediment exposures. *Anodonta kennerlyi* is a freshwater unionid mussel that was recently used for a caged mussel study on mercury uptake in a reservoir in BC (BC Research 1998).

A note of caution should be raised regarding the use of unionids for caged studies as part of EEM programs. Unionids have been used in a number of caged studies in which the endpoint of interest is bioaccumulation, whereas the focus for this technical guidance is survival, growth and condition. Obtaining sufficient numbers for use in a caged study may be very difficult and may jeopardize the status of natural unionid populations. Unfortunately, they are not commercially grown and are therefore not available in cultured form. Unionids are a long-lived group of organisms with slow recruitment. In the wild, they are often found as mature, large-sized, slow growing individuals and smaller, faster growing individuals are notoriously difficult to find.

As an alternative to unionids, sphaeriid clams might be considered suitable for assessing differences in the survival, growth, condition and reproductive rates of bivalves in freshwater receiving environments. Sphaeriid clams have short life cycles (generally one year), are abundant and widespread in distribution, and do not depend on fish hosts for a part of their life cycle (as unionids do). Sphaeriids are also generally more tolerant than unionids of low calcium concentrations in the environment (Stewart and Malley 1997), which may be an important factor in many regions of Canada. Sphaeriid clams are ovoviviparous (Mackie 1978a) and hold neonate clams in brood pouches on their gills until they are ready to be released (Mackie et al. 1974). Sphaeriid clams have been successfully used in the past to assess the effect of pulp mill effluent and other chemical stressors on growth and natality (Mackie 1978b; Sandusky and Sparks 1979; Stephenson and Mackie 1981). Potentially useful species for EEM studies include various genera of *Sphaerium* (particularly *S. striatinum*) and *Musculium* (particularly *M. securis*). The various taxa of the genus *Pisidium* may be too small for practical handling and, in addition, are taxonomically difficult for the non-specialist. Additional research may be required to demonstrate the utility of sphaeriid bivalves for use in EEM.

12.2.4.4 Source

Bivalves may be obtained either from natural populations or from culturing facilities. Where possible, bivalves should be obtained from a commercial grower who can give certification or other assurance that the supplied culture is single-species and taxonomy is confirmed. Animals obtained from a commercial facility have an environmental history that is well known and assurances of being uncontaminated are greater than for animals collected in the wild. In any study, some form of certificate should be provided to assure the identity of the species used (i.e., a letter from the supplier or a taxonomist's letter of validation).

All individuals used in a caged bivalve study should be from the same population, because different populations of the same species may differ in their response to the exposure. If wild populations are the only possible source, then they must be collected from an uncontaminated area. For wild populations, infaunal bivalves can be separated from sediment by hand or by gentle sieving. Care should be taken not to over-sample an existing natural population. If the species of choice uses byssal threads for attachment to substrate, a knife or scissors should be used to gently remove byssal threads from the outside of the shell so that the internal byssal glands are not damaged. Remove all epiphytic growth from bivalve shells gently by hand or by gentle use of a soft brush or scraper.

Collection permits for field-collected bivalves may be required by some local, provincial and federal agencies. Transplant or aquaculture permits may be required in advance from local and provincial agencies. In addition, if cages present a potential obstruction to navigation, then a permit may be required from Canadian Coast Guard and a Notice to Mariners may be required. The permitting process should be considered early in the study planning process, as it may take several weeks to months to process the necessary permits.

12.2.4.5 Size and Age

All bivalves used in a caged study should be from the same age class and as uniform as possible in size. Age class is more difficult to determine in wild specimens than in cultured specimens because wild populations may consist of several age classes, with individuals growing at different rates. Sub-adult (immature) specimens may be selected because they typically have the greatest potential for growth. On the other hand, it may be desirable to use adults that have just completed spawning so that growth during the exposure period reflects growth in both somatic and reproductive tissues. It is important to choose specimens that will not become reproductive and spawn during the exposure period because spawning may confound the results. Investigators should attempt to note the reproductive status of a subsample of the test specimens at the beginning and end of the test and note whether spawning occurred.

12.2.4.6 Number of Organisms

Power analysis should be used to determine the minimum number of bivalves to detect a specified "effect size". It is recommended that studies be designed to detect a 20 percent change in growth. Data sets to aid in predicting the number of required organisms may be available through Environment Canada.

The number of animals required per cage for the growth endpoint will depend on study design (e.g., number of cages per station; see more discussion on this in Section 12.2.7.1), species used, age of animals, variability in response to the station, and growth conditions at the station. The number of bivalves required to fill the cages will depend on study design in terms of :

- number of areas (e.g., exposure area plus number of reference areas);
- number of stations per area;
- number of moorings per station;
- number of cages per mooring (if the study is designed to address varying depths in the water column); and
- bivalves per cage.

In addition, if tissue samples are required for tissue analysis, then consideration must be given to the number of animals required to obtain a sufficient sample size for chemical analysis.

When the minimum number of animals required to achieve the desired power has been estimated, the number of bivalves collected needs to be determined, and should include:

- total number required for test set-up (see set of bullets above);
- plus the number required for baseline measurements at the beginning of the test;

- plus the number required to account for dead or dying individuals during transport and handling (dead or dying bivalves are identified by smell or if their shells do not close very quickly in response to tapping or do not close at all); and
- plus 50 to 100 percent more to account for size variability and allow you to obtain the size fraction you are targeting.

12.2.4.7 Handling and Holding Conditions

Salazar and Salazar (2000) provide detailed guidance on handling and holding conditions for bivalves, and this is summarized below. Test organisms should be handled as little as possible and should be deployed as soon as possible after collection. When handling is necessary, it should be done carefully, gently and quickly so that the bivalves are not unnecessarily stressed. Bivalves should be kept in well-aerated, clean-flowing water as long as possible between collection, sorting and deployment. If transporting bivalves for extended periods, keep them moist and cool by placing them in a cooler with frozen gel packs or wet ice (wet ice at the bottom of the cooler). Use newspaper or cloth towels to keep bivalves separated from the gel packs or wet ice.

For use of caged bivalves in EEM, there may be concerns about bivalves coming into contact with chemicals used in gel packs, chlorine present in wet ice, or other chemicals that may be released from newspaper and towels. The best strategy to avoid the need for holding at low temperature is to minimize time between collection and deployment. Bivalve supply should be as close to the deployment sites as possible. If there are extended periods (over one day) between collection and deployment, bivalves should be held in clean (uncontaminated), well-aerated water kept at the cool end of the seasonal environmental range for the test species. Placing small containers with bivalves within a larger container in which temperature is controlled should maintain bivalves at the desired temperature.

12.2.5 Study Design

Determining the appropriate study design is critical if the results are going to be meaningful. Study design for caged bivalves includes a number of components:

- sampling design;
- area and station selection;
- replication of cage stations and cages per station;
- timing and duration;
- modifying or confounding factors;
- supplementary measures;

- cage design; and
- mooring systems.

12.2.6 Sampling Design

There are five main sampling designs that are recommended for marine and freshwater assessments:

- Control-Impact (C-I) design;
- Multiple Control-Impact (MC-I) design;
- Simple Gradient (SG) design;
- Radial Gradient (RG) design; and
- Multiple Gradient (MG) design.

The C-I and MC-I designs are used to determine the magnitude of difference between homogeneous exposed and unexposed areas, while the SG, RG and MG designs examine changes in an effect along an effluent gradient. The C-I and MC-I designs are used when there are few qualitative levels of exposure. It is suggested that multiple reference areas be used rather than increasing sample sizes in one reference station (Environment Canada 1998). The SG, RG and MG designs may be used when there are many quantitative levels of exposure (Paine 2000) and also may be useful in discriminating among effects from sources other than pulp mill effluent. Guidance on selecting the appropriate design is provided in Sections 5.1.3.3 and 6.1.5.4 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Section 5.6 of the metal mining Guidance Document (Environment Canada 2001).

For example, in estuaries with complex tidal regimes and mixing regimes, an MC-I approach may be the most appropriate. A SG approach may be applicable to a river receiving environment where flow is unidirectional.

C-I and MC-I designs require some level of replication within the control and impact areas. Replication at each station may not be required for gradient designs, although an appropriate number of stations is required in order to discriminate between spatial patterns related to effluent discharge, and other spatial patterns in the environment. Replication is discussed below in Section 12.2.7. Ultimately, it is the responsibility of study designers to develop site-specific study designs that are scientifically defensible, robust, and suitably sensitive.

12.2.6.1 Area and Station Selection

Section 3 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Chapter 3, Section 4.2.2 and Section 5.8 of the metal mining Guidance Document (Environment Canada 2001) provides guidance for selection of multiple reference locations for a variety of receiving environments and is applicable to caged mussel studies. Reference areas should be as similar as possible to study areas in terms of:

- hydrodynamic conditions;
- temperature;
- salinity;
- dissolved oxygen; and
- food availability and quality.

Multiple reference stations may help to identify natural differences and variability among uncontaminated areas.

The exposure area is defined by plume characteristics. Plume delineation, as described in Section 2.2 of the pulp and paper Technical Guidance (Environment Canada 1998) and Section 3.3 in the metal mining Guidance Document (Environment Canada 2001), should provide sufficient information to model average effluent concentrations with distance from source and to identify the degree of vertical mixing in the water column. This information will assist in the selection of stations and depth of deployment in the water column. More guidance is provided below in Section 12.2.7.1.

Boating and recreational activities in exposure and reference areas should be considered during station selection to minimize the incidence of cage removal or damage due to boating activities and to minimize incidents of vandalism. Boat basins and areas with significant marine traffic may also result in exposure of bivalves to petroleum products, anti-foulants and other chemicals associated with boating activities.

12.2.7 Replication

12.2.7.1 Number of Cage Stations

For the exposure area, stations and depths for cage deployment may be chosen to represent a gradient of exposure to the effluent plume. The deployment stations of bivalve cages should be given careful consideration for a number of potentially interfering factors, such as.

- in estuarine and marine situations, effluent may be positively buoyant, resulting in a thin layer of "freshwater" effluent floating on denser saline water;

- in tidal situations, the mixing behaviour of effluent may be quite complex, resulting in low confidence about the average exposure concentration that bivalves may be exposed to. Tidal situations may require consideration of exposure area stations that are both "upstream" and "downstream" of the outfall; and
- in river discharge situations, effluent plumes may be quite long and narrow, meaning that there is little or no opportunity to meaningfully replicate cage stations within (i.e., across the long axis of) the plume.

For these and other reasons articulated below, it might be more meaningful to simply evaluate "distance from outfall" rather than effluent concentrations when designing caged bivalve studies. This approach would remain consistent with one of the objectives of EEM, which is to evaluate the magnitude and geographic extent of effects that may be related to the discharge of effluent.

12.2.7.2 Replication and Pseudoreplication

The number of replicate stations must be determined to address the sampling design and EEM objectives. The number of replicate stations and subsamples within replicate stations will be determined by power analysis. The allocation and distribution of replicate stations is dependent upon the sampling design. Guidance on the use of power analysis is provided in Section 6.1.5.5 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Sections 4.4.1.1 and 5.11 of the metal mining Guidance Document (Environment Canada 2001).

Any discussion of replication in a field study must consider the question of "pseudoreplication", defined by Hurlbert (1984) as "the use of inferential statistics to test for treatment effects with data from experiments where either treatments are not replicated (though samples may be) or replicates are not statistically independent". In ANOVA terminology, it is "the testing for treatment effects with an error term inappropriate to the hypothesis being tested" (see Box 1 for further discussion). See also Section 3 of the Technical Guidance Document (Environment Canada 1998) for further discussion on Hurlbert (1984) and the issue of pseudoreplication.

The issue of pseudoreplication is one that will continue to challenge field experiments for the foreseeable future. For EEM studies, where a single effluent stream is discharged into the receiving environment, it could be argued that no true replication is possible, and that any set of samples must be correlated with each other in ways that undermine inferential statistics. The onus will remain on study designers to develop approaches to EEM studies that minimize the risk of drawing incorrect conclusions as a result of the potentially "misleading" side-effects of pseudoreplication (see Table 1 in Hurlbert 1984).

It would appear that the best approach would be to evaluate (using analysis of variance (ANOVA) or other statistical techniques) whether there are significant differences between stations that happen to be strategically located around an effluent outfall. Any interesting correlation that may exist between significant differences and distance from the outfall (or the expected effluent concentration) may also be noted. It may be necessary to conduct further studies to attribute measured differences to the effluent.

This limitation is well recognized within the overall EEM context, and applies with equal force to FS and benthic invertebrate community survey.

Box 1 Discussion of Pseudoreplication in the Context of Hurbert (1984)

The discussion of pseudoreplication must refer back to one of the objectives of the EEM studies, which is to measure the *magnitude and geographicl extent* of any effects of effluent on specific aspects of the receiving environment and/or its biota. There is an important point implicit in this objective, and that is that effects we wish to detect are unique to the mill or mine being investigated. Every combination of mill or mine and receiving environment is unique, therefore care must be taken when generalizing to other receiving environments with respect to environmental effects associated with a single effluent. Also implicit in the objective of the EEM studies is that we wish to evaluate whether there is a statistically significant difference in the environment or biota caused by mill or mine effluent at various locations near the effluent outfall.

Returning to Hurlbert (1984), two classes of experiments are defined: *mensurative* (involving only the making of measurements at one or more points in space or time), and *manipulative* (involving two or more treatments, having as a goal the making of one or more comparisons, and organized so that different experimental units receive different treatments, and that the assignment of treatments to experimental units is or can be randomized). In the EEM context, we choose to regard locations as experimental units, rather than cages containing one or more bivalves. In making this decision, we regard the EEM studies to be mensurative in character, with the objective of assessing whether a biological property (such as survival or growth of bivalves) is different at one location than at another.

Hurlbert (1984) provides as illustrations several hypothetical examples of experiments. We refer specifically to Hurlberts's *Example 3*, where the decomposition of leaves contained in a number of mesh bags is being compared at two different locations (the 1 m and the 10 m isobath of a water body). Hurlbert notes (italics his):

"….out of laziness, we place all eight bags at a single spot on each isobath. It will still be legitimate to apply a significance test to the resultant data. However, and the point is the central one of this essay, if a significant difference is detected, this constitutes evidence only for a difference between the two (point) *locations*; one 'happens to be' a spot on the 1-m isobath, and the second 'happens to be' a spot on the 10-m isobath. Such a significant difference cannot legitimately be interpreted as demonstrating a difference between the two *isobaths*, i.e., as evidence of a 'treatment effect'. For all we know, such an observed significant difference is no greater than we would have found if the two sets of eight bags had been placed at two locations on the *same* isobath.

If we insist on interpreting a significant difference in Example 3 as a "treatment effect", then we are committing what I term *pseudoreplication*."

For more discussion about pseudoreplication, please consult Hurlbert (1984).

12.2.7.3 Applying Replication To Study Design

Several measurement endpoints are suggested for caged bivalve studies (survival; growth -change in length or wet weight; soft tissue fresh weight; condition). In order to assess these endpoints on a set of bivalves exposed at different locations in the field, many different configurations of cages and bivalves are possible. One possible approach is outlined below and illustrated in Box 2, although it is emphasized that site-specific conditions may make a wide variety of different options viable.

Assume sets of 100 bivalves are deployed at a reference area, and in the exposure area at distances of 50, 200 and 400 m away from an effluent outfall. A hypothetical data set was generated using random numbers, representing this scenario; the data can be considered to represent the soft tissue fresh weight of bivalves at the end of the exposure period (see Box 2). To generate this data set, four sets of 100 random numbers were generated, with the sets having expected mean values of 13 (reference), 10 (50 m), 11 (200 m) and 12 (400 m), and standard deviations of 2.0. Each set of 100 numbers was subsequently split into five groups of 20, representing bivalves deployed on individual moorings. Note that this approach assumes that moorings are replicates, and that the individual bivalves deployed on each mooring are sub-samples (not true replicates).

Three potential study designs are evaluated using these data. In the first design, all 100 bivalves at each Distance are considered to be replicates (although this is technically incorrect), and the resulting ANOVA has 3 degrees of freedom (DF) for Distance, and 396 DF for experimental error. The difficulty with this design is that only a single estimate of survival is available at each Distance, although tests of significance for the effect of area on growth are valid (ANOVA 1).

In the second design (ANOVA 2), bivalves are deployed in cages of 20 individuals, and each bivalve within each cage is treated as an individual sub-sample. In this analysis, five estimates of survival are available for each Distance, and the ANOVA (which separates out the Subsampling Error and considers the mean square error for "Cages Within Distance" to be the appropriate error term for testing effects of Distance) is more powerful than the analysis in ANOVA 1, despite being based on the same data.

In the third design (ANOVA 3), bivalves are again deployed in cages of 20 individuals, but for each cage, only the mean growth rate within cage is estimated. In this analysis, the test for effects of Distance has exactly the same power as the analysis using subsampling.

Box 2: Bivalve tissue dry weight as a function of distance from mill outfall (synthetic data)

50 m					200 m				
14.53	11.77	10.38	8.70	11.06	8.65	10.74	9.07	8.34	12.13
6.44	10.66	10.71	6.90	12.25	14.69	8.34	9.78	11.57	10.47
9.43	9.40	11.47	9.54	11.33	10.46	11.57	10.64	15.26	9.33
10.22	10.07	11.77	7.44	7.86	11.76	8.92	10.87	12.11	11.12
9.13	5.99	9.11	7.86	8.62	10.43	13.75	7.98	12.03	13.90
9.51	10.58	9.90	10.50	10.50	12.51	11.62	13.53	12.95	13.42
11.32	7.63	10.70	13.17	10.88	12.62	12.38	11.12	7.03	8.70
9.21	10.38	14.57	6.26	10.11	11.97	11.56	12.73	12.82	10.75
14.79	6.91	7.21	10.31	7.92	10.15	6.87	13.77	12.57	11.52
8.68	8.23	10.85	7.19	9.61	13.85	10.65	12.96	7.28	12.20
9.49	13.16	9.62	10.68	9.20	14.27	13.75	7.39	9.57	7.22
8.49	8.18	6.70	7.30	8.81	9.99	10.24	12.02	9.87	8.65
11.13	7.66	14.63	7.76	7.92	10.21	11.11	9.42	9.36	14.28
12.07	10.39	7.84	12.03	12.61	11.35	8.36	12.80	9.03	10.95
13.27	9.32	12.03	10.66	10.71	7.81	7.08	13.55	11.46	15.76
11.75	10.81	5.43	7.91	7.69	9.93	10.94	12.44	11.71	12.31
9.72	12.35	8.52	10.19	10.75	14.50	12.05	8.77	10.02	13.71
9.61	7.63	13.00	8.52	13.98	9.56	9.35	8.75	10.85	7.08
10.12	10.73	10.41	10.27	10.40	9.02	12.39	10.66	14.98	9.76
5.65	9.07	11.27	9.89	9.33	9.94	10.49	12.71	10.61	12.78
10.228	**9.546**	**10.306**	**9.154**	**10.077**	**11.184**	**10.608**	**11.048**	**10.971**	**11.302**

400 m					Reference				
11.02	11.01	8.69	10.90	9.93	17.18	15.75	7.76	15.34	11.59
13.22	10.99	9.91	12.00	13.77	13.22	13.72	12.84	13.87	10.75
12.30	13.98	13.31	11.14	12.09	11.77	12.47	12.29	11.99	10.54
9.10	10.32	9.48	11.36	12.38	14.58	14.02	9.73	11.85	14.51
12.12	12.30	12.34	14.42	11.12	13.11	16.24	12.80	12.50	15.52
11.90	12.00	9.28	8.51	11.80	13.60	12.65	11.64	12.37	11.31
14.33	12.13	12.15	9.40	11.57	15.27	11.78	13.77	10.25	14.00
10.32	9.19	11.82	14.16	10.65	13.67	14.43	14.40	12.45	10.99
13.10	9.43	12.12	12.07	11.57	15.34	13.64	15.73	16.51	11.67
13.16	12.90	11.84	11.72	9.79	8.81	15.83	15.91	9.54	15.05
9.36	11.04	10.31	13.93	9.65	17.26	12.52	10.75	10.13	10.53
13.51	10.75	11.15	12.32	13.45	13.53	14.02	16.12	12.56	13.04
13.96	14.54	10.50	8.41	8.03	9.82	12.42	15.16	12.42	13.94
11.81	11.93	15.36	14.90	12.81	11.55	15.70	14.04	9.90	12.86
14.24	9.25	14.85	10.15	10.16	15.61	12.16	11.25	17.93	10.17
11.82	10.83	13.74	9.80	13.18	11.17	9.41	11.29	13.74	13.99
16.00	12.65	13.19	10.83	13.10	10.14	15.03	12.54	11.92	13.81
12.42	10.32	9.45	12.74	8.68	9.95	13.21	9.47	16.02	12.18
11.38	14.03	10.97	13.90	13.24	12.39	8.95	15.58	11.13	11.15
11.88	15.04	14.50	13.28	14.18	15.81	11.38	10.07	15.75	11.17
12.348	**11.732**	**11.748**	**11.798**	**11.558**	**13.189**	**13.266**	**12.657**	**12.909**	**12.439**

ANOVA 1: Every bivalve considered a true replicate

Source	Sum of Squares	DF	Mean Square	F-ratio	P
Distance	492.34	3	164.11	39.46	<0.0001
Error	1647.04	396	4.16		

ANOVA 2: Cages considered to be replicates; individual bivalves considered to be subsamples

Source	Sum of Squares	DF	Mean Square	F-ratio	P
Distance	492.34	3	164.11	62.40	<0.0001
Cages Within Distance*	42.13	16	2.63		
Clams Within Cages (= Subsampling Error)	1604.91	380	4.22		

* Cages within distance is the appropriate error term in the design with subsamplin

ANOVA 3: Cages considered to be replicates; mean values for growth within cages used in the analysis

Source	Sum of Squares	DF	Mean Square	F-ratio	P
Distance	24.62	3	8.21	62.40	<0.0001
Error	2.11	16	0.13		

12 - 64

The point of this discussion is to demonstrate that for the EEM studies, it may be advantageous to deploy replicate cages containing multiple bivalves at each station, and to consider only the average performance within each cage. This approach confers the additional advantage of providing redundancy in case one or more cages are lost, and may simplify the construction and deployment of cages. So long as the statistical hypothesis testing is confined to evaluating whether there are significant differences between areas (without uniquely attributing effects to effluent exposure), the replication and statistical analysis is valid.

As general guidance, it is suggested that cages of 20 animals per cage be deployed for the survival endpoint and at least five cages be deployed at each station to evaluate the growth endpoint. These cages should be deployed on individual moorings, not five cages on one mooring, since the mooring is the most appropriate unit of statistical replication. However, practitioners should be encouraged to explore the power and robustness of potential study designs, using synthetic data (or empirical data where available) as an integral part of the study design process to determine the minimum number of animals required.

12.2.7.4 Timing and Duration

Timing of the studies should be such that:

- it coincides with a high growth period in natural populations so that growth is maximized and differences in growth rate among treatments may be more measurable; and
- it does not coincide with a spawning period if the test bivalves are adults.

The duration of exposure should be 60 to 90 days (see discussion in Salazar and Salazar 2000). This should provide sufficient time for effects to survival and growth to be manifested.

12.2.7.5 Modifying or Confounding Factors and Supplementary Measures

Results of caged bivalve studies will depend, at least in part, on natural factors such as temperature, food supply, other physical/chemical properties of the test environments, species selected, condition of test organisms, exposure method, and handling of test organisms. Exposure and reference areas should be as similar as possible with respect to the factors listed below to minimize confounding differences. It may be useful to measure some of the factors to assist in interpreting results. These factors are presented in brief below and reviewed in more detail in Section 6 of Salazar and Salazar (2000).

Life cycle of test organisms is critical to the success of the study. The most important consideration is spawning cycle because of possible interference with growth that can confound the interpretation of results. Tests should be conducted with individuals that are not likely to spawn during the exposure period. If immature bivalves are used, this would not be an issue.

Temperature of the receiving environment can affect chemical uptake, survival and growth. Water temperature may be monitored during the exposure period to determine possible influences on test results and to evaluate similarity in temperature between areas. Continuous temperature monitoring devices are inexpensive (approximately $100 to $120 each) and readily available. One monitoring device could be deployed at each station by attaching to a cage.

Lack of acclimation to deployment water quality conditions may cause interference. It may be necessary to acclimate test organisms gradually to deployment conditions, particularly if collection site conditions are markedly different from those at deployment areas.

Food supply is important because it affects survival, growth, condition and chemical uptake. Food availability may be sampled during the exposure period to aid in interpretation of results. Measurements could include one or more of the following (depending on feeding mechanism of the test species): chlorophyll a, particulate or organic carbon, suspended solids.

Current speed may have an effect on growth rate of bivalves. It is important that investigators have a general understanding of current patterns and velocities for each area and station to aid in interpreting results. Current meters may be considered for stations with simple flow regimes. Currents of stations with complex flow regimes might be difficult to assess using meters. In such situations, a qualitative understanding of currents is sufficient.

Salinity is of particular importance in estuarine areas, where salinity can range from 0 parts per thousand near the mouth of the river to 33 parts per thousand within a short distance. Salinity can also vary to the same extent at some locations during the tidal cycle. A species with a wide salinity tolerance range should be chosen for such situations, although there is no guarantee that survival will not be affected (see Grout and Levings 2000). Salinity at each station may be measured using an in situ monitor. Local conditions should be observed to determine if salinity fluctuations are likely.

Bivalve cages can become fouled with plant and animal growth in highly productive embayments. **Excessive fouling** can reduce or eliminate water flow through the cage material, resulting in highly stressful conditions to the test organisms. If such conditions are anticipated, the deployed cages should be examined for fouling during the exposure period, and fouling removed by hand or with the use of a stiff brush.

Variability in chemical concentrations in the water column as a result of intermittent discharges may make it difficult to estimate exposure concentrations. Monitoring of discharge volumes and concentrations of tracer chemicals in the discharge during the exposure period will help to estimate exposure and interpret results. Chemical concentrations in receiving water could be sampled, for example, at beginning and end of the test to confirm presence/absence of effluent and other chemical influences in the cage deployment areas. Guidance on tracers is provided in Section 2.2.2 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Section 4.4.1.1.2 of the metal mining Guidance Document (Environment Canada 2001).

12.2.7.6 Cage Designs

Cages with individual compartments are suggested so that each individual bivalve can be tracked through the study. The mesh size should be maximized to allow maximum water flow but small enough to contain the test animals. Individuals are assigned to compartments such that survival, growth and condition can be tracked on each bivalve.

A variety of cage designs are described in Salazar and Salazar (2000). A flat (i.e., two-dimensional) cage design, as shown in Appendix I, is suggested, as it is a convenient unit to work with a number of reasons (M. Salazar, pers. comm.), including:

- flat cages have a single layer of bivalves that is easy to handle during set-up, deployment and retrieval;
- flat cages can be easily adjusted to accommodate a variable number of bivalves; and
- flat cages are quickly and easily stacked in a minimal space during set-up and transit to the deployment sites.

PVC tubing is a convenient material to use for constructing cages. PVC should be water-supply grade and from a high quality source. PVC tubing should be soaked before use for at least 24 hours in flowing fresh or seawater to remove water-soluble and volatile chemicals. Alternative materials are described in Salazar and Salazar (2000: Section 9).

Individual test organisms are placed in the mesh bags and separated by using a plastic cable tie or other suitable tie. Sufficient space should be allowed in each compartment to permit test animals to grow during the exposure period.

Final cage dimensions depend on the size of the test organism and the number of organisms per cage. Typical mesh sizes for species like mussels and clams are 10 to 15 cm diameter with 5 mm mesh size. Each mesh bag should be long enough to accommodate the desired number of bivalves per bag plus sufficient material for attachment to the PVC frame. The length of the frame should be 5 cm longer than each mesh sock with bivalves, and the width should be 5 cm greater on each end than the adjacent mesh sock to the frame.

For freshwater clam deployments, a wide variety of cage designs have been used by investigators. It is the responsibility of the study designer to ensure that cage designs are appropriate to the test species and receiving environment.

12.2.7.7 Mooring Systems

The cages can be suspended in the water column by attaching them to mooring lines that have an anchor or weight on one end (e.g., iron chain links) and a surface or subsurface buoy attached to the other end. The following factors should be considered for deployment of cages (from Salazar and Salazar 2000):

- appropriate vessel to accommodate cages with associated mooring lines, floats and anchors;
- tidal height (to ensure cages are at the desired depth during both low and high tides);
- slope and nature of the bottom material (to ensure the cages do not slide down a steep slope during the exposure period); and
- boating activity and recreational activity in the vicinity of the cages.

A number of mooring configurations are described in Salazar and Salazar (2000: Section 9) and in the pilot studies (Applied Biomonitoring 2000; Andrews and Parker 1999). A mooring system for a cube-shaped cage (from Andrews and Parker 1999) is illustrated in Figure 6. It may be wise to include a retrieval line on each cage to ease recovery at the end of the exposure period.

12.2.8 Variables to be Measured

The test variables to be measured in caged bivalve studies for EEM studies are survival growth and condition, as described in more detail below. An example of a reporting format for recording survival and growth raw data and endpoints is provided in Table 9. Reproductive condition is not recommended at this time by the EEM Science Committee for the reasons outlined below.

The EEM Science Committee does not feel that, at this time, there are well-developed, standardized methods of evaluating the reproductive condition of bivalves. The EEM Science Committee feels that this is an area that requires further research and development before this type of endpoint can be incorporated into an EEM program. An example of reproductive condition is gonad index used by Mucklow (1996) on mussels (*Mytilus edulis* and *M. trossulus*) and also on mussels by the Gulf of Maine Council on the Marine Environment (GMCME 1992). Gonad index is calculated as the ratio between gonad dry weight and tissue dry weight (GMCME 1992).

Guidance can be found in Salazar and Salazar (2000: Section 11) on how to analyze bivalve tissues. Pulp and paper mills may be required to conduct a tissue analysis of dioxins and furans if:

a) the concentration of dioxins and furans in the effluent has exceeded the prescribed limits in *the Pulp and Paper Mill Effluent Chlorinated Dioxins and Furans*, or
b) effluent has contained measurable concentrations of 2,3,4,8-TCDD or of 2,3,7,8-TCDF within the meaning of *the Pulp and Paper Mill Effluent Chlorinated Dioxins and Furans*, or
c) an effect on fish tissue was reported in the most recent EEM cycle.

Metal mines may be required to conduct a tissue analysis of mercury if mercury has been detected in the effluent during effluent characterization. Other chemicals or metals (e.g. Cd, Pb, Se) in fish tissue may be analyzed, based on site specific concerns.

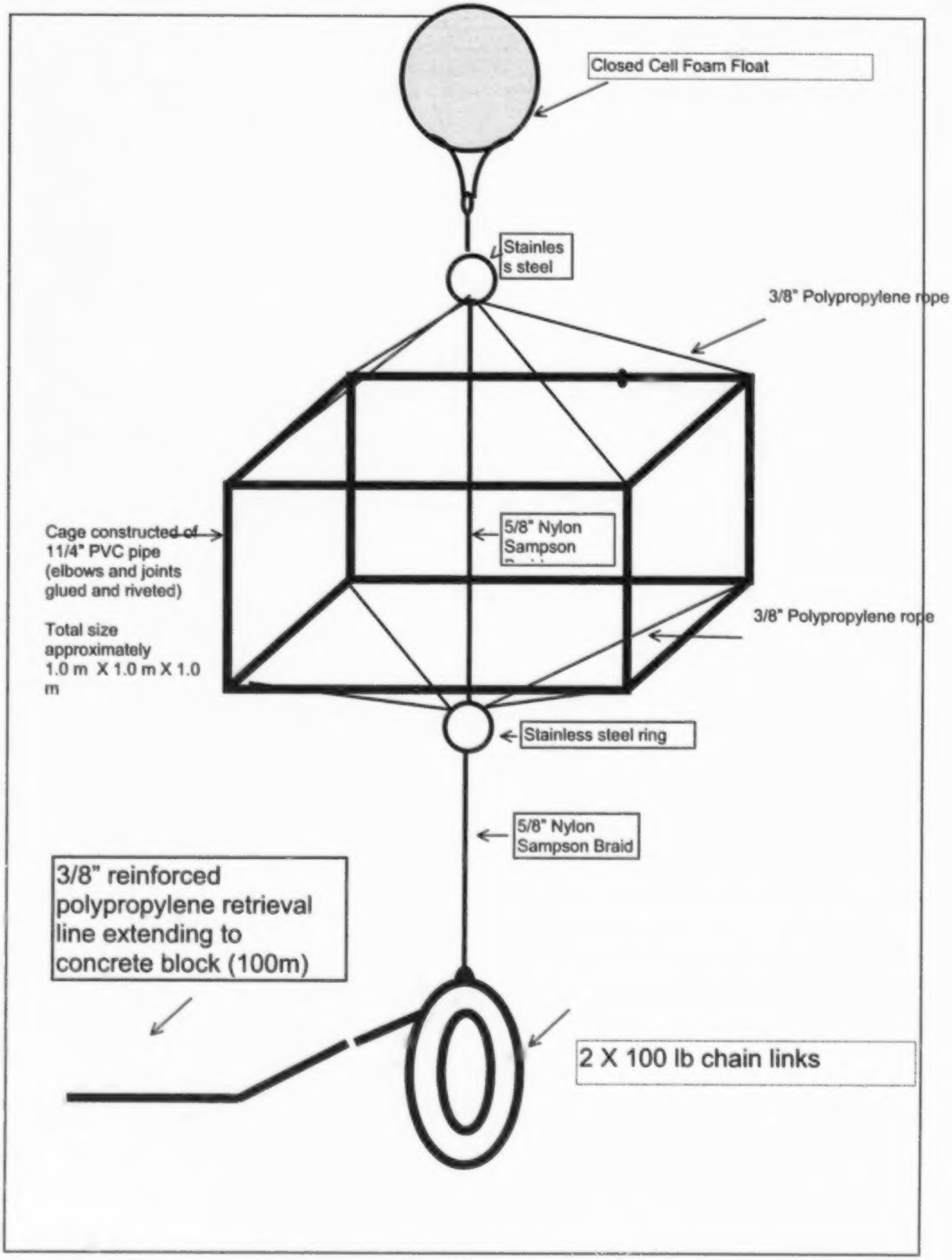

Closed Cell Foam Float
Stainles
s steel
3/8" Polypropylene rope
Cage constructed of 11/4" PVC pipe (elbows and joints glued and riveted)
Total size approximately 1.0 m X 1.0 m X 1.0 m
5/8" Nylon
Sampson
3/8" Polypropylene rope
Stainless steel ring
5/8" Nylon
Sampson Braid
3/8" reinforced polypropylene retrieval line extending to concrete block (100m)
2 X 100 lb chain links

Figure 12-6 Example of bivalve mooring design

Table 12-9 An Example of A Field Data Sheet For Recording Survival And Growth Raw Data

Caged Bivalve Field Data Sheet

Study Area ____________________ Cage Identifier ____________________
Coordinates: Latitude ____________N # of Mussels per cage
____
Longitude ____________W
Date/Time of Deployment ____________ Date/Time of Retrieval ____________
Days of Exposure ______ Depth of Water______ Depth of Cage ______

Mussel Identifier	Length (0.01 mm)	Length (0.01 mm)	Whole Animal Wet Weight (0.01 g)	Whole Animal Wet Weight (0.01 g)	Soft Tissue Fresh Weight (0.01 g)
	Start	End	Start	End	End
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					

Number of mussels alive at end of test ________ Percent Survival ________

Site Conditions:

Date	Time	Temperature (°C)	Dissolved Oxygen (mg/L)	Salinity (ppt)	Current Velocity (cm/sec)	Current Direction (degrees)

Field Staff: ________________________________

Notes:

____

12.2.8.1 Survival

Survival is not a particularly sensitive indicator of effects in caged bivalves, but is an important parameter to monitor. Survival can be easily determined and quantified, although it is possible to have some individuals missing at the test end due to shell decomposition. Bivalves are dead if they are gaping and do not close their shells when touched or tapped. Survival is expressed as percent of individual animals alive per cage at the end of the exposure period.

12.2.8.2 Growth

Growth is a measure of energy use and is a sensitive indicator of effects that is easy to measure. There are several types of growth measurements that should be made; measuring only one could provide misleading results. Growth measurements are outlined in Table 10.

Table 12-9 Growth Measurements

Growth Measurement	Measured at:		How Measured	Accuracy [2]
	Test Initiation	Test Termination [1]		
Whole animal wet weight (WAWW)	✓	✓	Analytical balance	0.01 g
Shell length	✓	✓	Vernier calipers; longest axis	0.01 mm
Soft tissue fresh weight	✓ (on sacrificed subsample)	✓	Analytical balance	0.01 g

[1] Test termination measurements should only be made on surviving individuals.

[2] Accuracy is based on larger bivalve species such as marine mussels and freshwater unionids. Accuracy may need to be adjusted downward for smaller-sized test species, such as sphaeriid clams.

Whole animal wet weight (WAWW) and shell length are non-destructive measurements and should be made on each individual at the beginning and end of the test. Soft tissue fresh weight should also be measured at the beginning and end of the test. However, as this is a destructive test, measurements at the beginning of the test should be made on a subsample of bivalves from the same population and size range as the test organisms. Soft tissue weight should be measured as fresh weight (e.g., tissue weight should be measured as soon as possible after test termination while the soft tissue is still fresh and has not been frozen or left to dry out). Weights (WAWW and soft tissue fresh weight) should be determined with an analytical balance with a measurement accuracy of 0.01g. Shell length should be determined with vernier calipers along the longest axis with a measurement accuracy of 0.01 mm. Please note that accuracy of these measurements may need to be adjusted downwards for smaller-sized test species.

Growth can be expressed in a number of ways:

- absolute growth: = absolute change in value from test initiation to test termination;
- growth rate: = absolute change in value per unit time, typically using one week as the time increment; or
- relative growth: = (final weight – initial weight) / initial weight; relative growth may be used when there is a significant difference among cages in initial weights; relative growth is expressed as a proportion and therefore an arcsin square root transformation of relative growth values may be appropriate prior to applying statistics. Green (1979: Page 52) provides useful advice on this transformation, noting that it is not usually required for proportion data in the range of 0.3 to 0.7, and that while it may not always help, it probably does no harm either.

Use the most appropriate expression of growth to suit the study design and site-specific characteristics.

1.6.3 Condition

Condition is a measure of how an animal stores its energy. There are more than one option that may be considered for calculating condition, as described below. Note that some of these methods require measurement of additional variables to those outlined in Section 12.2.8.2. The most appropriate method to calculate condition is left to the discretion of the investigators.

Weight (WAWW, dry shell or soft tissue weight) related to shell length

This is analagous to the Fulton Condition Index (Ricker 1975; Anderson and Neumann 1996) used in fisheries biology. This relationship may be characterized according to a conventional formula for a straight line (e.g., in Mackie and Flippance 1983), with slope (C) and intercept (b):

$\log W = b + (C \times \log L)$

High values of C imply that a bivalve has relatively high tissue weight at a given length, whereas low values may indicate that animals are not obtaining sufficient food, or are experiencing chronic stress that prevents them from thriving. This method of characterizing condition is suitable for assessing condition in wild bivalves where shell length is expected to be quite variable. However, since animals used for caged bivalve studies are screened for uniform length at test initiation, this method will not be reliable. This method may also not be suitable for bivalves because shell length and tissue weight are influenced by different factors in the environment (M. Salazar, pers. comm.).

Soft tissue weight related to shell weight

This method of characterizing condition uses soft tissue weight and shell weight. An ANOVA can be conducted or, more simply, soft tissue weight can be divided by shell weight. Grout and Levings (2000) measured condition of blue mussels as the ratio of tissue weight to shell weight. They found that condition distinguished caged mussels in a high survival zone (condition index 1.10 to 1.42) from caged mussels in a low survival zone (condition index 0.82 to 0.96).

Soft tissue weight related to shell volume

This method was used by Mucklow (1996; based on Seed 1968) to calculate condition by dividing soft tissue dry weight by shell volume, measured as length x width x height. In a study on wild blue mussel populations, Mucklow (1996) concluded that seasonal patterns in condition index were variable and influenced by a number of natural factors, including food availability and physiological energy requirements.

12.2.9 Methods for Test Initiation, Cage Deployment and Retrieval, Test Termination

These methods are described in detail in Salazar and Salazar (2000: Section 11) and summarized below.

12.2.9.1 Test Initiation

The first step is to sort all bivalves into the desired size range(s). By selecting bivalves within a narrow size/age range, the investigator can be relatively confident that the individuals will have similar growth potential. Determining the age of wild-collected bivalves may be difficult if not impossible for some species; length is therefore used to obtain individuals with similar growth rates. Commercial growers can often provide bivalves of a known age.

The size range selected for test organisms depends on the test species, test organisms supply, and target age. Test organisms should be selected within a narrow size range. Some studies have selected bivalves with a narrow size range and randomly allocated them among cages (e.g., Andrews and Parker 1999). Other studies have used multiple size classes and allocated bivalves to cages such that there is no significant difference among cages in terms of initial size and weight (see Salazar and Salazar 2000: Section 11).

Test organisms should be kept cool and moist during the sorting stage to prevent stress, damage or death (see Section 12.2.4.7). Shell lengths and WAWW are recorded for each animal and the animals assigned to individual holding containers (e.g., ice cube tray) that are labelled according to specific location within the cage. This allows measurements on each individual test organism to be repeated at test termination. Animals in a sacrificed subsample (statistically appropriate number) of test organisms will be measured for shell length, WAWW, tissue weight and shell weight.

For quality control purposes, approximately 10 percent of the measurements should be repeated and recorded by a different investigator. It is best to use an electronic spreadsheet for recording all measurements to reduce transcription errors. It is possible to electronically link measurement tools (e.g., vernier calipers and analytical balance) with a spreadsheet to provide an additional quality control measure.

All test organisms needed to fill the cages, plus a subsample for destructive sampling (tissue weight), should be sorted and measured prior to distribution to socks and cages. Test organisms are then distributed to individual mussel socks before they are attached to cages. An example of a labeling and distribution scheme is provided in Salazar and Salazar (2000: Section 11).

To ensure statistical similarity in WAWW and shell length among stations, an ANOVA can be conducted. This can be programmed into a spreadsheet such that the data are analyzed by cage and by station (if more than one cage is to be deployed at each station). If statistically significant differences are found, the test animals should be redistributed to eliminate this difference. This may require opening mesh socks and re-closing with a plastic tie.

12.2.9.2 Cage Deployment

Ensure that an appropriate vessel is chartered to ensure safe transport and deployment of all cages, moorings and floats. Moorings and floats may be attached to the cages on the vessel on route to site. When on site, cage location should be identified using Global Positioning System (GPS), or some other reliable method (e.g., nearby on-shore reference points). During the exposure period, the cages may be inspected (e.g., using SCUBA divers) for presence, damage and fouling.

12.2.9.3 Cage Retrieval and Test Termination

Cages should be retrieved using various location identifiers (e.g., GPS, depth sounders) and the use of a grappling hook if a retrieval rope was used for each cage.

Details on test termination are provided in Salazar and Salazar (2000: Section 11) and outlined here. At the processing site, bivalves from all bags for each cage should be processed together. It is essential that the order and orientation of each bivalve is maintained during all of the end-of-test measurements. Individuals can be removed from mesh socks into labelled individual containers (e.g., ice cube trays) to facilitate measurement.

At 5 to 10 minutes prior to making length and weight measurements, set the tray(s) into a tub containing clean water. Individuals that float indicate that air is trapped in between the valves of the shell. When all animals in the tray(s) have opened their valves slightly and are not longer floating (5 to 10 minutes), length and weight measurements may begin. Start with WAWW and shell length measurements, then carefully shuck each individual and blot the tissue before making tissue fresh weight measurements. Note that if tissues are to be analyzed for chemical parameters, then care must be taken not to introduce any contaminants from the blotting material.

12.2.10 Data Analyses

Data analyses and interpretation of the results should be appropriate for the study design. Caged mussel studies for EEM are designed to determine if there are significant effects on biota in the vicinity of effluent outfalls. This can be accomplished by using either a control-impact (= reference-exposure) or a gradient (= regression) type of design, as discussed in Section 1.4.1. Statistical procedures appropriate for each effect endpoint (i.e., survival, growth, and condition) are summarized in Table 11

Table 12-10 Summary of Statistical Procedures to Analyze Effect Endpoints for Caged Bivalve Studies

Type of Response	Effect Endpoint	Statistical Procedures
Survival	Survival	ANOVA
Energy Use	Growth – based on shell length	ANOVA (regression analysis for gradient designs)
	Growth – based on WAWW	ANOVA (regression analysis for gradient designs)
	Growth – based on soft tissue fresh weight	ANOVA (regression analysis for gradient designs)
Energy Storage	Condition	ANOVA or ANCOVA
Bioaccumulation*	Tissue Chemistry*	Multivariate statistics

* May not be required for EEM (determined on a site specific basis), but may be used to assist in data interpretation.

The first step is to generate summary statistics for each parameter (i.e., WAWW, shell length, soft tissue fresh weight) on a cage and station basis.

The second step is to determine whether there are significant differences among replicate cages for each of the parameters measured before deployment (if they were not measured post-distribution and before deployment). This involves assessing the data for normality and homogeneity of variances.

The final step is to use the appropriate statistical test (see Table 11) for the study design. In general, ANOVA and multiple comparison tests are used for hypothesis testing and comparison among stations. For C-I and MC-I designs using ANOVA and Analysis of Co-Variance (ANCOVA) procedures, detailed guidance is provided in Sections 5.1.5.3 to 5.1.5.5 and in 5.1.6.1 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Section 5.16 of the metal mining Guidance Document (Environment Canada 2001). If statistical differences are found, a multiple range test, or its non-parametric counterpart, can be used to determine which stations are different from the others. Linear and multiple regression analyses (using variables regressed on distance) may generally be used to establish relationships among variables along exposure gradients. For ANCOVA analyses of condition, the covariates will be dictated by the condition formula that is used.

12.2.11 Tissue Concentrations

12.2.11.1 Dioxins and Furans

If a mill is required to measure dioxins and furans in fish tissue, then measuring these chemicals in caged bivalves may be considered. However, there are some precautions that should be taken with this approach, including the following:

- duration of exposure should be chosen such that dioxins and furans are likely to be accumulated to detectable levels in bivalve tissue;
- consider whether bivalves would be harvested (commercially or recreationally) for human consumption in the area; and
- ensure that there are sufficient number of bivalves included in the design to obtain enough tissue for analysis.

There may be mills for which caged bivalves would not be considered suitable for measuring uptake of dioxins and furans. Any proposal to use caged bivalves for dioxin and furan measurements would require approval by the Regional Authorization Officer.

12.2.11.2 Mercury

Section to be completed

12.2.11.3 Other Chemicals

If there is a requirement for a particular mill or mine to measure tissue concentrations to assess bioavailability of particular chemicals, then refer to guidance in Salazar and Salazar (2000: Section 11) on sampling design considerations and handling.

12.2.12 Reporting

Quality assurance and quality control (QA/QC) considerations for caged bivalve studies should follow those outlined for the Fish Survey in Section 5.1.8 of the pulp and paper Technical Guidance Document (Environment Canada 1998) and Section 4.14 of the metal mining Guidance Document (Environment Canada 2001). QA/QC measures apply to the following components of caged bivalve studies:

- study design;
- field sampling;
- sample processing/laboratory analysis;
- data analysis; and
- reporting.

12.2.13 References

Anderson, R.O. and R.M. Neumann. 1996. Length, weight, and associated structural indices. Pp. 447-482. In: B.R. Murphy and D.W. Willis (eds.). Fisheries Techniques. Second Edition. American Fisheries Society, Bethesda, Maryland.

Andrews, S. and R. Parker. 1999. Draft Report (II). An Environmental Quality Evalutaion of Pictou Harbour, Nova Scotia, Using Caged Bivalves; *Mytilus edulis*. Prepared by Environmental Protection Branch of Environment Canada (Dartmouth, NS) for The Pictou Harbour Environmental Protection Project, New Glasgow, NS.

ASTM 2001. Standard Guide for Conducting In-situ Field Bioassays with Marine, Estuarine and Freshwater Bivalves, E-2122. American Society forTesting and Materials. Conshohocken, PA, pp. 1546-1575.

BC Research Inc. 1998. Methyl Mercury Impact Assessment for Nechako Reservoir Submerged Timber Salvage Operations: Caged Mussel Study. Prepared for Ministry of Environment, Lands and Parks, Environmental Protection – Skeena Region, Environmental Impact Section, Smithers, BC.

Beckvar, N., S. Salazar, M. Salazar and K. Finkelstein. 2000. An in situ assessment of mercury contamination in the Sudbury River, Massachusetts, using transplanted freshwater mussels (*Elliptio complanata*). Can. J. Fish. Aquat. Sci. 57: 1103-1112.

Clarke, A.H. 1973. The freshwater molluscs of the Canadian interior basin. Malacologia 13, (1-2): 1-509.

Clarke, A.H. 1981. The freshwater molluscs of Canada. National Museum of Natural Sciences, National Museums of Canada, Ottawa, ON. 446 pp.

Couillard, Y., P.G.C. Campbell, A. Tessier, J. Pellerin-Massicotte and J.C. Auclair. 1995a. Field transplantation of a freshwater bivalve, *Pyganodon grandis*, across a metal contamination gradient. I. Temporal changes in metallothionein and metal (Cd, Cu, and Zn) concentrations in soft tissues. Can. J. Fish. Aquat. Sci., 52: 690-702.

Couillard, Y., P.G.C. Campbell, A. Tessier, J. Pellerin-Massicotte and J.C. Auclair. 1995b. Field transplantation of a freshwater bivalve, *Pyganodon grandis*, across a metal contamination gradient. II. Metallothionein response to Cd and Zn exposure, evidence for cytotoxicity, and links to effects at higher levels of biological organization. Can. J. Fish. Aquat. Sci., 52: 703-715.

Courtenay, S.C., W.R. Parker and G.P. Rawn. 1998. Proceedings of a workshop to assess alternatives to the fish survey component of the environmental effects monitoring program for Canadian pulp and paper mills. Can. Tech. Rep. Fish. Aquat. Sci., 2233: viii + 108 pp.

Day, K.E., J.L. Metcalfe and S.P. Batchelor. 1990. Changes in intracellular free amino acids in tissues of the caged mussel, *Elliptio complanata*, exposed to contaminated environments. Arch. Environ. Contam. Toxicol., 19: 816-827.

Environment Canada. 2001. Metal Mining Guidance Document for Aquatic Environmental Effects Monitoring. Environment Canada. Hull, Québec (draft).

Environment Canada. 1998. Pulp and Paper Technical Guidance for Aquatic Environmental Effects Monitoring EEM/1998/1.

Freeman, K.R., K.L. Perry, T.G. DiBacco and D.J. Scarratt. 1994. Observations on two Mytilid species from a Nova Scotian mussel farm. Can. Tech. Rep. Fish. Aquat. Sci. No. 1969.

GMCME (Gulf of Maine Council on the Marine Environment). 1992. Evaluation of Gulfwatch 1991 Pilot Project of the Gulf of Maine Marine Environmental Monitoring Plan. Gulf of Maine Council on the Marine Environment, State Planning Office, August, ME.

Goldberg, E.D., V.T. Bowen, J.W. Farrington, G. Harvey, J.H. Martin, P.L. Parker, R.W. Risebrough, W. Robertson, E. Schneider and E. Gamble. 1978. The Mussel Watch. Environ. Conserv., 5(2): 101-125.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley & Sons, Toronto, ON. 257 pp.

Grout, J.A. and C.D. Levings. 2000. Effects of acid mine drainage from an abandoned copper mine, Britannia Mines, Howe Sound, British Columbia, Canada, on transplanted blue mussels (*Mytilus edulis*). Marine Environmental Research, 51(3): 265-288.

Hinch, S.G. and R.H. Green. 1989. The effects of source and destination on growth and metal uptake in freshwater clams reciprocally transplanted among south central Ontario lakes. Can. J. Zool., 67: 855-863.

Hornbach, D.J., T.E. Wissing and A.J. Burky. 1982. Life-history characteristics of a stream population of the freshwater clam *Sphaerium striatinum* Lamarck (Bivalvia: Pisidiidae). Can. J. Zool., 60: 249-260.

Hurlbert, S.H. 1984. Pseudoreplcation and the design of ecological field experiments. Ecol. Monog., 54: 187-211.

Jernelov, A. 1996. The International Mussel Watch: A global assessment of environmental levels of chemical contaminants. Sci. Tot. Environ., 188 (Suppl. 1): 37-44.

Mackie, G.L. 1978a. Effects of pollutants on natality of *Musculium securis* (Bivalvia: Pisidiidae). The Nautilus, 92 (1): 25-33.

Mackie, G.L. 1978b. Are sphaeriid clams ovoviviparous or viviparous? The Nautilus, 92: 145-147.

Mackie, G.L., S.U. Qadri and A.H. Clarke. 1974. Development of brood sacs in *Musculium securis* Bivalvia: Sphaeriidae. The Nautilus, 88: 109-111.

Mackie, G.L. and L.A. Flippance. 1983. Growth dynamics in *Sphaerium rhomboideum* (Bivalvia: Pisidiidae). Can. J. Zool., 61: 868-873.

Malley, D.F., A.R. Stewart and B.D. Hall. 1996. Uptake of methyl mercury by the floater mussel, *Pyganodon grandis* (Bivalvia, Unionidae) caged in a flooded wetland. Environ. Toxicol. Chem., 6: 928-936.

McMahon, R.F. 1991. Mollusca: Bivalvia. In J.H. T.G. Horp and A.P. Covich (eds). Ecology and Classification of North American Freshwater Invertebrates. Academic Press Inc., San Diego, California. Pp. 315-399.

Metcalfe-Smith, J.L., R.H. Green and L.C. Grapentine. 1996. Influence of biological factors on concentrations of metals in the tissues of freshwater mussels (*Elliptio* *complanata* and *Lampsilis radiata radiata*) from the St. Lawrence River. Can. J. Fish. Aquat. Sci., 53: 205-19.

Mucklow, L.C. 1996. Effects of season and species on physiological condition and contaminant burdens in mussels (*Mytilus edulis L.* and *Mytilus trossulus G.*): implications for mussel watch programs. Thesis submitted in partial fulfillment of requirements for Master of Environmental Studies, School for Resource and Environmental Studies, Dalhousie University, Halifax, NS.

Newell, R.I.E. 1989. Species Profiles: Life Histories and Environmental Requirements of Coastal Fishes and Invertebrates (North and Mid-Atlantic) – Blue Mussel. U.S. Fish Wildl. Serv. Biol. Rep. 82 (11.102). US. Army Corps of Engineers, TR EL-82-4.

Nichols, S.J. 2000. Nonindigenous Species Research and Outreach. Life History and Ecological Requirements of the Zebra Mussel – North American Experience Through 1992. http://www.nsgo.seagrant.org/research/nonindigenous /zmlifehistory.html.

O'Connor. 1992. Recent trends in coastal environmental quality: Results from the first five years of the NOAA Mussel Watch Project. Washington, DC. U.S. Department of Commerce, National Oceanic and Atmospheric Administration (NOAA), National Ocean Service.

Paine, M.D. 2000. EEM Design. Short-Course presented at the 27th Annual Aquatic Toxicity Workshop, October 1-4th, St. John's, NF.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin 191. Department of the Environment and Fisheries and Marine Service. Ottawa, ON.

Salazar, M. 2001. Applied Biomonitoring, Seattle, WA. Personal Communication via conference call, March 2001.

Salazar, M.H. and S.M. Salazar. 1995. In situ bioassays using transplanted mussels: I. Estimating chemical exposure and bioeffects with bioaccumulation and growth. Pp: 216-241. In: J.S. Hughes, G.R. Biddinger and E. Mones (eds). Environmental Toxicology and Risk Assessment – Third Volume. American Society for Testing and Materials (ASTM), Philadephia, PA.

Salazar, M. and S. Salazar. 2000. Standard Guide for Conducting In-Situ Field Bioassays with Caged Marine, Estuarine and Freshwater Bivalves. (Draft dated November 17, 2000). American Society for Testing and Materials (ASTM), Philadelphia, PA. Accepted for publication in 2001.

Sandusky, M.J. and R.E. Sparks. 1979. Investigations of declines in fingernail clam (Musculium transversum) populations in the Illinois River and Pool 19 of the Mississippi River. Bull. Amer. Malacol. Union Inc., 1979: 11-16.

Seed, R. 1968. Factors influencing shell shape in *Mytilus edulis* L. J. Mar. Biol. Ass. U.K., 48: 561-584.

Skidmore, D. and K.K. Chew. 1985. Mussel aquaculture in Puget Sound. Washington Sea Grant Publication, Seattle, Washington.

Stephenson, M. and G.L. Mackie. 1981. Effects of 2, 4-D treatment on growth and survival of fingernail clams (Bivalvia: Pisidiidae) in artificial pond ecosystems. Bull. Amer. Malacological Union, 1981: 19-22.

Stewart, P.L. 1994. Environmental requirements of the blue mussel (*Mytilus edulis*) in Eastern Canada and its response to human impacts. Can. Tech. Rep. Fisheries Aquat. Sci., 2004. 41 pp.

Stewart, R. and D.F. Malley. 1997. Technical evaluation of molluscs as a biomonitoring tool for the Canadian mining industry. Aquatic Effects Technology Evaluation (AETE) Program Project 2.3.1. Sponsored by Canada Center for Mineral and Energy Technology (CANMET) and the Mining Association of Canada (MAC).

Virginia Tech. 2000. Fish and Wildlife Information Exchange, Conservation Management Institute, Species Profile. Blacksburg, VA. (**Error! Bookmark not defined.**, Virginia Tech.

Waldock, M.J., J.E. Thain and M.E. Waite. 1996. An assessment of the value of shell thickening in *Crassostrea gigas* as an indicator of exposure to tributyltin. Pp. 219-237. In: M.A. Champ and P. F. Seligman (eds). Tributyltin: Environmental Fate and Effects. Chapman and Hall, London.

Williams, J.D., M.L. Warren Jr., K.S. Cummings, J.L. Harris and R.J. Neves. 1993. Conservation status of freshwater mussels of the United States and Canada. Fisheries, 18 (9): 6-22.